

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CÂMPUS CORNÉLIO PROCÓPIO
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**

ALANA REGINA BIAGI SILVA LISBOA

**APLICAÇÃO DE *CLUSTERING* E MÉTRICAS À ANÁLISE DE *LOG*
PARA AVALIAÇÃO AUTOMÁTICA DE USABILIDADE DE
APLICAÇÕES INTERNET RICAS**

DISSERTAÇÃO DE MESTRADO

CORNÉLIO PROCÓPIO

2015

ALANA REGINA BIAGI SILVA LISBOA

**APLICAÇÃO DE *CLUSTERING* E MÉTRICAS À ANÁLISE DE *LOG*
PARA AVALIAÇÃO AUTOMÁTICA DE USABILIDADE DE
APLICAÇÕES INTERNET RICAS**

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Tecnológica Federal do Paraná – Câmpus Cornélio Procópio como requisito parcial para a obtenção do título de Mestre em Informática

Orientador: Prof. Dr. Luciano Tadeu Esteves Pansanato

CORNÉLIO PROCÓPIO

2015

Dados Internacionais de Catalogação na Publicação

- L769 Lisboa, Alana Regina Biagi Silva
Aplicação de *clustering* e métricas à análise de *LOG* para avaliação automática de usabilidade de aplicações internet ricas / Alana Regina Biagi Silva Lisboa. – 2015.
102 f. : il. ; 30 cm
- Orientador : Luciano Tadeu Esteves Pansanato .
Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Informática. Cornélio Procópio, 2015.
Referências: p. 94 – 98.
1. Sites da web – Programação – Programas de computador. 2. Medição de software. 3. Software de aplicação – Desenvolvimento. 4. Software - Testes. 5. Informática – Dissertações. I. Pansanato, Luciano Tadeu Esteves, orient. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Informática. III. Título.

CDD (22. ed.) 004



Título da Dissertação Nº 05:

“Aplicação de *Clustering* e Métricas à Análise de Log para Avaliação Automática de Usabilidade de Aplicações Internet Ricas”.

por

Alana Regina Biagi Silva Lisboa

Orientador: **Prof. Dr. Luciano Tadeu Esteves Pansanato**

Esta dissertação foi apresentada como requisito parcial à obtenção do grau de MESTRE EM INFORMÁTICA – Área de Concentração: Computação Aplicada, pelo Programa de Pós-Graduação em Informática – PPGI – da Universidade Tecnológica Federal do Paraná – UTFPR – Câmpus Cornélio Procópio, às 15h30 do dia 22 de maio de 2015. O trabalho foi APROVADO pela Banca Examinadora, composta pelos professores:

Prof. Dr. Luciano Tadeu Esteves Pansanato
(Presidente)

Prof. Dr. André Pimenta Freire
(UFLA-MG)

Prof. Dr. Marco Aurélio Graciotto Silva
(UTFPR-CM)

Visto da coordenação:

Carlos Nascimento Silla Junior
Coordenador do Programa de Pós-Graduação em Informática
UTFPR Câmpus Cornélio Procópio

A Folha de Aprovação assinada encontra-se na Coordenação do Programa.

Dedico este trabalho à minha família.

AGRADECIMENTOS

A Deus, por me dar a oportunidade de estudar.

Ao meu orientador Prof. Dr. Luciano Tadeu Esteves Pansanato, pela competência, tranquilidade e ensinamentos transmitidos durante todo o curso.

Ao meu esposo Júnior, pela compreensão e apoio.

Aos meus filhos, Léo e Helo, por entenderem que tudo que fiz foi por eles, e que esse tempo que eu deixei de dedicar a minha família será recompensado.

Aos meus pais, pelo incentivo.

A minha irmã Lyvia e meu cunhado Arthur, pelas dicas e orientações.

Ao colega de trabalho Adriano Rivolli por todo o trabalho que teve com a ferramenta WAUTT.

As minhas amigas da Coordenadoria de Gestão de Recursos Humanos, por se empenharem ainda mais no trabalho nos momentos que eu precisava.

A todos os profissionais da Coordenadoria de Gestão de Tecnologia da Informação, por todas as vezes que prontamente se disponibilizaram a colaborar no que foi preciso.

A todos, que de alguma forma, me ajudaram a vencer essa etapa.

Nada há encoberto que não venha a ser revelado, nem oculto, que não venha a ser conhecido (Matheus 10:26b)

RESUMO

LISBOA, Alana Regina Biagi Silva Lisboa. **Aplicação de *Clustering* e Métricas à Análise de Log para Avaliação Automática de Usabilidade de Aplicações Internet Ricas**. 2015. 102 f. Dissertação (Mestrado em Informática) – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná.

O desenvolvimento de sistemas com eficácia, eficiência e satisfação dos usuários tem feito com que a usabilidade se torne uma característica importante na avaliação de qualidade de um produto de *software*. A análise da interação do usuário com o sistema é uma das formas de mensurar essa característica. Novas tecnologias permitem a criação de sistemas com foco na qualidade de interação do usuário, com grande parte do processamento ocorrendo no lado cliente, como as Aplicações Internet Ricas ou simplesmente RIAs, de *Rich Internet Applications*. No entanto, os dados de interação do usuário armazenados no servidor web são insuficientes para extrair conhecimento útil sobre a qualidade da interação. Este trabalho apresenta uma abordagem para avaliação automática de usabilidade através da aplicação de *clustering* e métricas na análise de *log* de RIAs. A abordagem proposta utiliza a *Web Application Usage Tracking Tool*, WAUTT, para capturar a interação do usuário e o algoritmo X-Means para realizar o agrupamento dos dados. O resultado da avaliação automática emprega métricas que fornecem informação quantitativa a respeito da usabilidade de sistemas web e que poderão ser utilizadas pelo avaliador para auxiliar na tomada de decisão. As informações obtidas com as métricas foram comparadas a um método de avaliação tradicional, a avaliação heurística, para corroborar os resultados obtidos pela abordagem proposta. O *clustering* foi uma técnica útil para reduzir o volume de dados e permitira ao avaliador concentrar os seus esforços em determinados grupos de usuários com comportamentos semelhantes. A aplicação de métricas aos *clusters* permitiu realizar uma avaliação de usabilidade quantitativa para auxiliar a tomada de decisão em relação a algumas das subcaracterísticas de usabilidade. A avaliação automática de usabilidade realizada tem potencial para auxiliar no desenvolvimento de sistemas web.

Palavras-chave: Avaliação de Usabilidade. Análise de Log. *Clustering*. Métricas. Aplicações Internet Ricas.

ABSTRACT

LISBOA, Alana Regina Biagi Silva Lisboa. **Application of Clustering and Metrics to the Log Analysis for Automatic Usability Evaluation of Rich Internet Applications.** 2015. 102 f. Dissertação (Mestrado em Informática) – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná.

The development of systems with effectiveness, efficiency and user satisfaction has made that usability become an important feature in quality assessment of software products. The analysis of user interaction data with the system is one way to measure this feature. New technologies allow the creation of systems with focus on quality of user interaction and a large part of the processing occurs on the client side, as Rich Internet Applications, from RIAs for short. However, user interaction data stored on the webserver are insufficient to extract useful knowledge about the quality of interaction. This dissertation presents an approach for automatic usability evaluation by applying clustering and metrics in log analysis of RIAs. The proposed approach uses the Web Application Usage Tracking Tool, WAUTT, to capture user interaction and the X-means algorithm to perform clustering. The result of the automatic evaluation applies metrics that provide quantitative information about the usability of web systems and may be used to assist evaluators in making decisions. Information obtained from metrics were compared to a traditional evaluation method, heuristic evaluation, to corroborate the results obtained by the proposed approach. Clustering was a useful technique to reduce the volume of data and allowed the evaluator focus their efforts on certain groups of users with similar behaviors. The application of metrics enabled perform a quantitative usability evaluation to assist the decision making process about some usability subcharacteristics. The automatic usability evaluation conducted as the proposed approach has the potential to assist in developing web systems.

Keywords: Usability Evaluation. Log Analysis. Clustering, Metrics. Rich Internet Applications.

LISTA DE SIGLAS

CSV	<i>Comma Separated Values</i>
DOM	<i>Document Object Model</i>
GOMS	<i>Goals, Operators, Methods and Selection Rules</i>
IEC	<i>International Electrotechnical Commission</i>
IP	<i>Internet Protocol</i>
ISO	<i>International Organization for Standardization</i>
JSON	<i>JavaScript Object Notation</i>
OCS	<i>Open Conference System</i>
RIA	<i>Rich Internet Applications</i>
SAM	<i>Sequence Alignment Method</i>
SQuaRE	<i>System and Software Quality Requirements and Evaluation</i>
WAUTT	<i>Web Application User Tracking Tool</i>
XML	<i>Extensible Markup Language</i>
UI	<i>User Interface</i>
URL	<i>Uniform Resource Locator</i>

LISTA DE QUADROS

Quadro 1 – As Dez Heurísticas de Usabilidade de Nielsen	25
Quadro 2 – Trabalhos Relacionados.....	41
Quadro 3 – Características das Ferramentas.....	44
Quadro 4 – Comparação entre as Avaliações do DREDS.....	73
Quadro 5 – Comparação entre as Avaliações do SEAC.....	74

LISTA DE TABELAS

Tabela 1 - Níveis de Severidade dos Problemas de Usabilidade.....	26
Tabela 2 – Padronização dos dados para execução do agrupamento.....	50
Tabela 3 – Execução do X-Means, Tarefa 1, Sistema DREDS	56
Tabela 4 – Métricas do cluster 0, Tarefa 1, Sistema DREDS.....	57
Tabela 5- Métricas do cluster 1, Tarefa 1, Sistema DREDS.....	57
Tabela 6 – Métricas do cluster 3, tarefa 1, Sistema DREDS.....	57
Tabela 7 – Execução do X-Means, tarefa 2, Sistema DREDS.....	58
Tabela 8 – Métricas do cluster 0, Tarefa 2, Sistema DREDS.....	58
Tabela 9 – Métricas do cluster 1, Tarefa 2, Sistema DREDS.....	58
Tabela 10 – Execução do X-Means, Tarefa 3, Sistema DREDS.....	59
Tabela 11 – Métricas do cluster 0, Tarefa 3, Sistema DREDS.....	59
Tabela 12 – Métricas do cluster 1, Tarefa 3, Sistema DREDS.....	60
Tabela 13 – Métricas do cluster 2, Tarefa 3, Sistema DREDS.....	60
Tabela 14 – Execução do X-Means, Tarefa 4, Sistema DREDS.....	61
Tabela 15 – Métricas do cluster 0, Tarefa 4, Sistema DREDS.....	61
Tabela 16 – Métricas do cluster 2, Tarefa 4, Sistema DREDS.....	61
Tabela 17 – Métricas do cluster 3, Tarefa 4, Sistema DREDS.....	62
Tabela 18 – Execução do X-Means, Tarefa 5, Sistema DREDS.....	62
Tabela 19 – Métricas do cluster 0, Tarefa 5, Sistema DREDS.....	62
Tabela 20 – Métricas do cluster 1, Tarefa 5, Sistema DREDS.....	63
Tabela 21 – Síntese das métricas calculadas, Sistema DREDS.....	64
Tabela 22 – Execução do X-Means , Sistema SEAC.....	67
Tabela 23 – Métricas do cluster 1, Sistema SEAC.....	68
Tabela 24 – Métricas do cluster 2, Sistema SEAC.....	68
Tabela 25 – Síntese das métricas calculadas, Sistema SEAC.....	69
Tabela 26 – Quantidade de problemas encontrados na avaliação heurística do DREDS.....	71
Tabela 27 – Quantidade de problemas encontrados na avaliação heurística do SEAC.....	72

LISTA DE FIGURAS

Figura 1 – Exemplo de entrada em arquivo de <i>log</i> de servidor.....	23
Figura 2 – Divisões da Série SQuaRE.....	29
Figura 3 – Características de Qualidade de Produto de <i>Software</i>	30
Figura 4 – Características de Qualidade em Uso.....	31
Figura 5 – Modelo de Referência/ Avaliação de Qualidade de Produto de <i>Software</i>	32
Figura 6 – Modelo de Referência para Mensurar a Qualidade do Produto de <i>Software</i>	33
Figura 7 – Interface administrativa para acesso aos dados capturados pela WAUTT.....	35
Figura 8 – Estrutura do Projeto de Avaliação de Usabilidade Proposto.....	49
Figura 9 – Exemplo do arquivo de <i>log</i> de eventos gerado pela ferramenta WAUTT.....	51
Figura 10 – Exemplo do arquivo de <i>log</i> de eventos com dados numéricos.....	51
Figura 11 – Página do Sistema DREDS.....	54
Figura 12 – Página de acesso ao SEAC.....	66

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 JUSTIFICATIVA	16
1.2 OBJETIVOS.....	17
1.3 ORGANIZAÇÃO DA DISSERTAÇÃO	18
2 FUNDAMENTAÇÃO TEÓRICA.....	19
2.1 USABILIDADE	19
2.2 AVALIAÇÃO DE USABILIDADE	21
2.2.1 Análise de <i>Log</i>	22
2.2.2 Avaliação Heurística	24
2.3 AVALIAÇÃO AUTOMÁTICA DE USABILIDADE	26
2.4 PADRÕES PARA AVALIAÇÃO DA QUALIDADE DE <i>SOFTWARE</i>	29
2.5 CAPTURA DA INTERAÇÃO DO USUÁRIO	34
2.6 MINERAÇÃO DE DADOS.....	36
3 TRABALHOS RELACIONADOS.....	39
3.1 ANÁLISE DE <i>LOG</i> PARA AVALIAÇÃO DE USABILIDADE.....	39
3.2 ANÁLISE E COMPARAÇÃO	41
4 UM PROCESSO DE AVALIAÇÃO AUTOMÁTICA DE USABILIDADE	46
4.1 DEFINIÇÕES.....	46
4.2 ETAPAS DO PROCESSO	47
4.2.1 Estabelecer os requisitos de avaliação	47
4.2.2 Especificar a avaliação.....	48
4.2.3 Projetar a avaliação	48
4.2.4 Executar a avaliação.....	49
4.2.5 Concluir a avaliação.....	52
5 AVALIAÇÃO AUTOMÁTICA DE USABILIDADE USANDO <i>CLUSTERING</i> E MÉTRICAS.....	53
5.1 AVALIAÇÃO DO SISTEMA DREDS	53
5.2 SÍNTESE DOS RESULTADOS: DREDS	63
5.3 AVALIAÇÃO DO SISTEMA SEAC	65
5.4 SÍNTESE DOS RESULTADOS: SEAC.....	68
6 COMPARAÇÃO ENTRE AVALIAÇÃO AUTOMÁTICA DE USABILIDADE E AVALIAÇÃO HEURÍSTICA	70
6.1 AVALIAÇÃO HEURÍSTICA DO SISTEMA DREDS	70
6.2 AVALIAÇÃO HEURÍSTICA DO SISTEMA SEAC	71
6.3 COMPARAÇÃO: SISTEMA DREDS.....	72
6.4 COMPARAÇÃO: SISTEMA SEAC	74
6.5 LIMITAÇÕES DA ABORDAGEM	75
7 CONCLUSÕES.....	76
7.1 CONTRIBUIÇÕES	77
7.2 TRABALHOS FUTUROS	77

REFERÊNCIAS	79
APÊNDICE A - Atividades Propostas para Avaliação Automática - DREDS.....	84
APÊNDICE B - Atividades Propostas para Avaliação Automática - DREDS.....	90
APÊNDICE C - Atividade Proposta para Avaliação Automática - SEAC.....	93
APÊNDICE D - Interfaces do SEAC.....	94
APÊNDICE E - Resultado da Avaliação Heurística - DREDS.....	100
APÊNDICE F - Resultado da Avaliação Heurística - SEAC.....	103

1 INTRODUÇÃO

Com a crescente disseminação da Internet e o aumento da demanda por sistemas web, a exigência dos usuários em relação à qualidade desses sistemas tem crescido consideravelmente (RIVERO; CONTE, 2011). Um sistema web com baixa qualidade tem grandes chances de ser preterido pelo usuário, considerando a grande quantidade de aplicações com objetivos semelhantes disponíveis na Internet (MACK; NIELSEN, 1994; NIELSEN; LORANGER, 2006).

A usabilidade é a capacidade de um produto ser usado por usuários específicos para atingir objetivos específicos com eficácia, eficiência e satisfação em um contexto específico de uso (INTERNATIONAL..., 1991). Esse conceito está relacionado à qualidade da interação dos usuários com um *software* e depende de vários aspectos, como a facilidade de aprender, a eficiência e a satisfação do usuário (FREITAS; DUTRA, 2009). A definição de métodos e técnicas para avaliar a usabilidade tem sido um dos investimentos da indústria de desenvolvimento de *software* e um dos objetivos da pesquisa em Engenharia de Web (*Web Engineering*) (CONTE et al., 2007).

Diversas variáveis determinam o custo da avaliação de usabilidade: o escopo (quanto mais elementos a serem avaliados, maior o custo e maior o tempo para coletar e analisar os dados), o equipamento utilizado (quanto mais sofisticado o equipamento, maior o custo) e o pessoal (quanto maior o número de participantes e avaliadores envolvidos, maior o custo e o tempo) (MINISTÉRIO..., 2010). Com o objetivo de reduzir o custo da avaliação de usabilidade podem ser conduzidos testes remotos e automáticos de usabilidade, com usuários e avaliadores separados no tempo e espaço (MACK; NIELSEN, 1994). A automação da avaliação da usabilidade pode, portanto, ser uma forma promissora para aperfeiçoar as abordagens existentes (BASTIEN, 2010; IVORY; HEARST, 2001).

Uma das maneiras de avaliar a usabilidade de um sistema de forma remota e automática é através de teste com usuários (*user testing*). Os métodos de teste com usuários são os mais utilizados dentre as outras abordagens (FERNANDES et al., 2011). Em especial, a análise de *log* é um tipo de teste com usuários em que são analisadas as interações do usuário registradas em arquivos de *log* gerados durante a utilização de um sistema web (FERNANDES et al., 2011; IVORY; HEARST, 2001; MACK; NIELSEN, 1994; PAGANELLI; PATERNÒ, 2002).

O uso de arquivos de *log* que são armazenados no servidor web para a avaliação de usabilidade é uma estratégia conveniente e confiável (JANSEN, 2006). No entanto, os

arquivos de *log* de servidor têm a limitação de conter somente a informação sobre as páginas (URLs) visitadas pelo usuário e não oferecer dados detalhados sobre a interação real do usuário com um sistema web.

Nos sistemas web modernos, usualmente chamados de Aplicações Internet Ricas (de *Rich Internet Applications* – RIAs) (RIVOLLI et al., 2014), a interface pode ser uma única página que inclui subpáginas sob demanda que gerenciam toda a interação do usuário, de maneira semelhante a uma aplicação *desktop*. Esse paradigma evita a atualização de toda a página a cada interação e permite aos sistemas carregar, apresentar e alterar de maneira independente os elementos da página. Assim, os dados sobre a interação do usuário que podem ser armazenados no servidor web não são suficientes para obter informação detalhada sobre o uso real do sistema. No caso de RIAs, o interesse é trabalhar com arquivos de *log* com informação no nível de elementos da página e de eventos associados com a interação do usuário.

1.1 JUSTIFICATIVA

Os arquivos de *log* da interação do usuário com um sistema web que utiliza o paradigma RIA podem ser muito extensos, tornando inviável a análise manual em uma avaliação de usabilidade. Portanto, o emprego de processos automáticos é necessário para facilitar o trabalho dos profissionais envolvidos (avaliadores). Neste trabalho é apresentada uma aplicação de *clustering* e métricas para análise de *log* voltadas a avaliação automática de usabilidade de sistemas web.

A importância da aplicação de *clustering* está no fato de que o avaliador pode priorizar seus esforços apenas no registro da interação dos usuários que representem um determinado padrão de comportamento, por exemplo, considerando apenas os mais frequentes ou os que ocorrem de maneira isolada. Os registros da interação dos demais usuários podem ainda ser avaliados em outro momento, dependendo da disponibilidade de recursos e restrições de custo.

A importância da aplicação de métricas está na obtenção de indicadores quantitativos como resultado da avaliação automática de usabilidade. Esses indicadores podem auxiliar a tomada de decisão, através da análise de dados numéricos ao invés de somente realizar procedimento subjetivo, que depende única e exclusivamente da visão e experiência de um ou mais avaliadores.

1.2 OBJETIVOS

O objetivo geral desse trabalho foi apresentar uma abordagem para a avaliação automática de usabilidade de Aplicações Internet Ricas com base na aplicação de *clustering* e métricas à análise de *log*. Os objetivos específicos foram os seguintes:

- Definir uma abordagem de avaliação automática de usabilidade baseada somente na análise de *log*, sem realizar comparações entre o ideal e o que foi realizado pelo usuário.

O foco deste objetivo é não utilizar modelos de tarefa para comparação entre os passos realizados pelo usuário com aqueles considerados como ideais para a conclusão de uma tarefa. Esse tipo de abordagem torna a avaliação mais flexível, podendo ser aplicada em diversos contextos sem a necessidade de intervenção do avaliador na etapa inicial do processo.

- Executar a técnica de *clustering* em arquivos de *log* com dados da interação do usuário com RIAs.

Este objetivo consiste no uso de informação no nível de elementos da página e de eventos associados com a interação do usuário com RIAs como entrada para o algoritmo de *clustering*. Como muitos dos dados obtidos com essa interação não são numéricos, existe a necessidade de uma etapa prévia de pré-processamento para a posterior execução de *clustering*.

- Empregar métricas que forneçam indicadores de usabilidade para utilização pelo avaliador como apoio à tomada de decisão.

Este objetivo está centrado na apresentação de indicadores com informações menos subjetivas do que análise realizada por avaliadores com apresentação de relatórios exclusivamente descritivos. A participação do avaliador na etapa inicial seria por meio da definição dos valores considerados críticos para os indicadores calculados a partir das métricas.

- Comparar a proposta de avaliação automática de usabilidade com uma avaliação de usabilidade tradicional.

Este objetivo consiste na comparação dos resultados obtidos com a aplicação de *clustering* e métricas à análise de *log* com os resultados da avaliação heurística. Os dois métodos de avaliação seriam realizados em uma mesma RIA para posterior comparação dos resultados.

1.3 ORGANIZAÇÃO DA DISSERTAÇÃO

No Capítulo 2 é apresentada a fundamentação teórica sobre os principais conceitos relacionados à usabilidade, avaliação de usabilidade, avaliação automática de usabilidade, padrões para avaliação da qualidade de *software*, captura da interação do usuário e à mineração de dados. No Capítulo 3 são descritos os trabalhos relacionados à análise de *log* para avaliação de usabilidade. No Capítulo 4 é apresentado um processo de avaliação automática de usabilidade definido com base nos padrões para avaliação da qualidade de *software*. No Capítulo 5 é apresentada a abordagem proposta de avaliação automática de usabilidade aplicada a dois sistemas web. No Capítulo 6 é apresentada a comparação dos resultados da abordagem proposta com a avaliação heurística realizada nos dois sistemas web. No Capítulo 7 são apresentadas as conclusões, com foco nas contribuições e nos trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica está dividida em seis seções. Na primeira parte, apresentada na Seção 2.1, a usabilidade é descrita no contexto das normas ISO/IEC. A Seção 2.2 detalha o processo de avaliação de usabilidade e os dois métodos utilizados neste trabalho: a avaliação heurística e a análise de *log*. A Seção 2.3 apresenta um panorama da avaliação automática de usabilidade. Os padrões que definem os requisitos de qualidade de *software* e a forma de avaliação da qualidade, descritos no conjunto de normas da ISO/IEC 25000 – Série SQuaRE, são apresentados na Seção 2.4. A série SQuaRE define como avaliar as características de usabilidade de um *software* empregando técnicas de modo organizado e sistemático. As Seções 2.5 e 2.6 descrevem, respectivamente, a maneira como os dados que servirão de base para avaliação de usabilidade são capturados e a forma de mineração desses dados.

2.1 USABILIDADE

O termo usabilidade foi utilizado pela primeira vez como substituto da expressão “*user-friendly*” no início da década de 80. Um sistema “*user-friendly*” era considerado um sistema amigável, fácil de usar e operacionalizar. No entanto, essa expressão era considerada excessivamente vaga e subjetiva. A partir de então, surgiram diferentes abordagens para definir a usabilidade (DIAS, 2007):

- Definições orientadas ao produto: ligadas às características ergonômicas;
- Definições orientadas ao usuário: associadas ao esforço ou atitude do usuário;
- Definições baseadas no desempenho do usuário: ligadas à forma de interação do usuário, com foco na facilidade de uso e no grau de aceitação dos produtos;
- Definições orientadas ao contexto de uso: associadas à execução de tarefas específicas num ambiente determinado.

Em 1991, a Norma ISO/IEC 9126 (INTERNATIONAL..., 1991) descreveu a usabilidade em uma abordagem orientada ao produto. Usabilidade era considerada um conjunto de atributos de *software* relacionado ao esforço necessário para utilizá-lo e ao julgamento individual desse uso por um conjunto de usuários. Com a reedição dessa mesma norma, em 1998, as necessidades do usuário foram incluídas no conceito de usabilidade, que passou a ser definido pelas características de qualidade de *software*: funcionalidade,

confiabilidade, eficiência, possibilidade de manutenção e portabilidade (INTERNATIONAL..., 1998).

Em uma definição mais detalhada, a norma ISO/IEC 25010 (INTERNATIONAL..., 2011) estabeleceu o conceito de usabilidade considerando as características de qualidade de *software* e o ponto de vista do usuário e seu contexto de uso, como a capacidade de um produto ser usado por usuários específicos para atingir objetivos com eficácia, eficiência e satisfação em um contexto específico de uso. Para o contexto deste trabalho, essa foi a definição adotada para usabilidade.

O contexto de uso é a combinação dos usuários, tarefas, equipamentos e ambiente em que o produto é usado. A usabilidade pode ser considerada como a qualidade da interação entre o usuário e o sistema e depende de características tanto do sistema quanto do próprio usuário. Assim, a usabilidade está diretamente associada ao contexto operacional e aos diferentes tipos de usuários, tarefas e ambientes. A alteração em qualquer um desses aspectos pode alterar a usabilidade de um sistema (DIAS, 2007).

A eficácia está relacionada à precisão e à completeza com que os usuários atingem objetivos específicos. A precisão está associada à correspondência entre a qualidade do resultado e o critério especificado (para avaliação), enquanto a completeza é a proporção do quanto esse objetivo foi alcançado. A eficácia mede a adequação da interface, fornecendo indicativos relacionados à possibilidade de concluir as tarefas propostas de forma correta (PREECE et al., 2012).

A eficiência é uma característica quantitativa e diz respeito ao tempo total e aos recursos necessários para realizar uma tarefa específica ou à quantidade de erros ocorridos na execução dessa tarefa (INTERNATIONAL..., 1998). O sistema deve ser eficiente a ponto de permitir que o usuário atinja altos níveis de produtividade nas suas atividades. Essa característica de qualidade está intimamente ligada à experiência dos usuários. Para avaliar a eficiência de uso de um sistema é comum estabelecer um determinado nível de experiência desejado dos usuários e medir o tempo gasto por uma amostra desses usuários para a realização de tarefas típicas (DIAS, 2007).

A satisfação do usuário refere-se a percepções, sentimentos e opiniões dos usuários a respeito de um sistema, normalmente identificados por questionários ou entrevistas (DIAS, 2007). Um sistema eficiente, que gera resultados eficazes, pode não ser aceitável se a reação do usuário ao sistema for negligenciada.

O desempenho é outra métrica de usabilidade tão importante quanto a satisfação do usuário (DIAS, 2007). O desempenho reúne aspectos de eficácia e eficiência na sua definição:

habilidade do usuário em realizar tarefas para as quais o sistema foi desenvolvido e no seu contexto de uso.

O objetivo de se desenvolver sistemas considerando as características de usabilidade é permitir uma interação com qualidade. O foco na usabilidade deve proporcionar a criação de interfaces transparentes de maneira a não dificultar o processo, permitindo ao usuário pleno controle do ambiente sem se tornar um obstáculo durante a interação.

2.2 AVALIAÇÃO DE USABILIDADE

Avaliação de usabilidade é o processo sistemático de coleta de dados responsável por obter informações sobre como um ou mais usuários devem utilizar um sistema para uma determinada tarefa em certo tipo de ambiente (PREECE et al., 2012). O objetivo é verificar o desempenho (eficácia e eficiência) da interação do usuário e obter indícios da sua satisfação, identificando problemas durante a realização de tarefas específicas em seu contexto de uso (DIAS, 2007).

Os métodos de avaliação de usabilidade são divididos em três grandes grupos: métodos de inspeção, métodos de teste com usuários e métodos baseados em modelos (DIAS, 2007).

Os métodos de inspeção, também conhecidos como métodos analíticos ou de prognóstico, são caracterizados pela não participação direta dos usuários. Normalmente, os avaliadores que usam esses métodos se baseiam em regras ou recomendações estabelecidas previamente para identificar os problemas de usabilidade. São exemplos de métodos de inspeção para avaliação de usabilidade: inspeção de usabilidade formal, inspeção ou percurso pluralístico, inspeção de componentes, inspeção de consistência, inspeção ou percurso cognitivo, inspeção baseada em padrões, inspeção baseada em guias de recomendações e guias de estilo e avaliação heurística.

Ao contrário dos métodos de inspeção, os métodos de teste com usuários são caracterizados pela participação direta dos usuários. Os testes podem ser realizados por meio de questionários e entrevistas ou através de técnicas de observação e monitoramento do sistema em situações reais. São classificados como métodos de teste com usuários: entrevistas, questionários, testes empíricos, análise de *log* e *think aloud*.

Os métodos baseados em modelos, chamados de métodos de modelagem analítica, têm o objetivo de prever a usabilidade de um sistema a partir de modelos cognitivos ou

representações de sua interface e/ou de seus usuários. A intenção desses métodos é representar a interação do usuário e como ficam as características relacionadas ao entendimento, conhecimento, pretensões ou reações. São exemplos de métodos baseados em modelos os métodos da família GOMS (*Goals, Operators, Methods and Selection Rules*). Esses métodos empregam técnicas de modelagem e de análise de tarefas para prever o desempenho do usuário com determinada interface e os problemas de usabilidade encontrados.

Neste trabalho, a análise de *log*, um método de teste com usuários, é utilizada para a avaliação automática de usabilidade de aplicações interativas web. Para verificar a validade da abordagem proposta para a avaliação automática de usabilidade, a avaliação heurística, um método de inspeção, foi realizada com as mesmas aplicações e os resultados obtidos foram utilizados para realizar uma comparação. Esses dois métodos, análise de *log* e avaliação heurística, são descritos nas subseções seguintes.

2.2.1 Análise de *Log*

A análise de *log* é um método de avaliação de usabilidade classificado como método de teste com usuários, que pode ser realizada local ou remotamente. Esse método consiste no emprego de técnicas para analisar os dados referentes à interação do usuário com o sistema que, geralmente, ficam registrados em arquivos de *log* armazenados em servidores (CYBIS et al., 2010; IVORY; HEARST, 2001).

A possibilidade de ser realizada remotamente é uma das vantagens da análise de *log*. Dessa maneira, os usuários estão em seus ambientes naturais, o que minimiza a ocorrência dos problemas de interação artificiais, como aqueles gerados quando os testes são realizados em laboratório, quando o usuário percebe que está sendo monitorado e se sente constrangido. Além disso, como os usuários estão separados no tempo e espaço, é desnecessário realizar a avaliação em tempo real, o que permite envolver uma grande quantidade de usuários com baixo custo (MACK; NIELSEN, 1994).

O objetivo da análise de *log* é verificar os registros de todas as interações realizadas entre o usuário e o sistema para descobrir padrões de comportamentos dos usuários que possam indicar problemas de usabilidade. O uso dessa técnica gera dados relacionados a padrões de uso, usabilidade de produtos, informações sobre o acesso a determinada parte do

sistema, conteúdo mais acessado, caminhos realizados pelo usuário, dentre outros (PREECE et al., 2012).

De forma geral, cada registro em um arquivo de *log* de servidor contém: endereço IP que faz a requisição, data e horário da requisição, URL requisitada, tamanho do arquivo requisitado, protocolo empregado na requisição, navegador e sistema operacional usado pelo computador requisitante. A Figura 1 representa uma transação realizada com sucesso (código 0200), contendo o endereço IP da máquina requisitante (10.20.32.83), a data e horário do acesso, o tipo de requisição efetuada (solicitação da página index.html) e o tamanho do arquivo que retornou ao cliente (304).

```
10.20.32.83 [27/Oct/2013:11:10:34 -0200] "GET /index.html HTTP/1.1" 200 304
```

Figura 1 – Exemplo de entrada em arquivo de *log* de servidor
Fonte: Autoria própria.

As informações obtidas a partir de um arquivo de *log* de servidor podem ser organizadas em quatro tipos (CYBIS et al., 2010), apresentadas resumidamente pela lista a seguir:

1. Visitas: número de usuários, duração de cada acesso por usuário, quantidade de páginas acessadas em um determinado período, local de onde partiu a requisição do usuário, tipo do navegador e sistema operacional utilizado.
2. Referência: lista de endereços web de onde partem os usuários, termos de pesquisa que foram digitados para acessar o sistema.
3. Conteúdo: as páginas mais visitadas.
4. Navegação: os caminhos percorridos pelos usuários para alcançar o objetivo.

Os dados constantes em um arquivo de *log* são recursos interessantes para análise de usabilidade, considerando que os usuários estão em seus ambientes naturais, realizando tarefas reais. No entanto, trazem informações bastante superficiais em relação aos objetivos da análise de usabilidade, registrando apenas o caminho percorrido pelos usuários e a duração dessas visitas (CYBIS et al., 2010; OLINER et al., 2012; RIVOLLI et al., 2014).

Em RIAs, os dados que podem ser armazenados no servidor sobre a interação do usuário não são suficientes para extrair informação detalhada sobre o uso real da aplicação. Um sistema web tradicional consiste em múltiplas páginas, atualizadas a cada interação do usuário. Em sistemas com tecnologia RIA, pode haver uma única página com várias

subpáginas que gerenciam todas as interações com usuário. Essa tecnologia evita que páginas inteiras sejam atualizadas a cada interação e permite que apenas alguns elementos sejam atualizados. No entanto, é necessário que a interação dos usuários com todos os elementos da interface seja capturada para obter dados detalhados que possam servir de base para uma avaliação de usabilidade.

Neste trabalho, a ferramenta utilizada para realizar a captura da interação do usuário com o sistema web avaliado foi a *Web Application Usage Tracking Tool* (WAUTT) (RIVOLLI et al., 2014), que além de capturar as informações existentes nos arquivos de *log* de servidor, também permite registrar as interações do usuário ocorridas em sistemas web que utilizam tecnologia RIA.

2.2.2 Avaliação Heurística

A avaliação heurística é um método de inspeção para avaliação de usabilidade de sistemas interativos realizada por especialistas. O objetivo é identificar problemas de usabilidade que serão corrigidos ao longo do processo de desenvolvimento do sistema (NIELSEN, 1994). A avaliação heurística é o método de avaliação de usabilidade mais utilizado, uma vez que é fácil de ser aplicado e tem baixo custo (BLANDFORD *et al.*, 2004; KAROULIS; POMBORTSIS, 2003; HOLLINGSED, NOVICK, 2007; ALVES et al., 2014). Além disso, é considerado um método eficiente (NIELSEN, 1994), com capacidade de adaptação a diversos domínios de aplicação (PREECE *et al.*, 2012), sendo adequado também para sistemas web (RUSU et al., 2011; MACK; NIELSEN, 1994).

Na avaliação heurística, os avaliadores analisam o fluxo de interação exigido para completar uma tarefa real e fazem o julgamento dos elementos interativos da interface do sistema baseando-se em heurísticas, princípios ou padrões de usabilidade. Heurística é uma parte do conhecimento capaz de sugerir ações plausíveis a seguir ou ações implausíveis a evitar (LENAT, 1982). Assim, heurísticas são regras informais de julgamento que guiam as pessoas numa rápida tomada de decisão. Essas heurísticas servem como sugestões de como projetar uma boa interface e podem ser definidas pelos avaliadores ou especialistas.

O número de avaliadores depende da análise do contexto de uso do sistema e do custo benefício. A recomendação é que a avaliação seja realizada por três a cinco especialistas na área. As avaliações realizadas por especialistas podem produzir ótimos resultados em relação à rapidez da avaliação e à quantidade e importância dos problemas de usabilidade

identificados. No entanto, esse resultado está diretamente relacionado à competência dos avaliadores e às estratégias de avaliação empregadas. Pesquisas realizadas por Nielsen (1994) fornecem indicações sobre a quantidade de avaliadores para possibilitar a identificação de problemas de usabilidade. Cinco avaliadores especialistas são capazes de identificar 95% dos problemas de usabilidade de um sistema, enquanto que o mesmo número de avaliadores novatos identifica apenas 50% dos problemas (CYBIS et al., 2010).

Para evitar a influência de um avaliador na análise de outro avaliador, uma vez que esse processo é um julgamento de valor, cada um dos especialistas analisa a usabilidade separadamente, inspecionando e verificando todos os elementos várias vezes e realizando comparações com as heurísticas escolhidas. No entanto, não há restrição para que o avaliador considere outras heurísticas que julgue relevante.

As heurísticas consideradas neste trabalho foram as definidas por Nielsen (1994) baseadas em 294 tipos de erros encontrados usualmente em estudos empíricos. O Quadro 1 apresenta um resumo dessas heurísticas.

Nº	Heurística
1	Visibilidade do estado atual do sistema: O usuário precisa de <i>feedback</i> informando continuamente o que está acontecendo e dentro de um tempo razoável.
2	Correlação entre o sistema e o mundo real: A informação deve ser contextualizada; o sistema deve evitar termos técnicos, utilizando a linguagem do usuário e em ordem lógica e coerente.
3	Controle e liberdade do usuário: O usuário deve conseguir desfazer ou refazer uma operação realizada incorretamente.
4	Consistência e padrões: Funções, palavras ou situações iguais devem se identificadas da mesma maneira; a mesma ação deve ser apresentada na mesma localização e da mesma forma para facilitar o reconhecimento.
5	Prevenção de erros: Um sistema deve ser projetado para que não seja necessário enviar uma mensagem de erro; a prevenção de falhas deve ser configurada.
6	Reconhecimento ao invés de memorização: A interface do sistema deve oferecer ajuda contextual e informações capazes de orientar as ações do usuário; o usuário não deve ser obrigado a memorizar informações.
7	Flexibilidade e eficiência de uso: O sistema deve permitir ao usuário personalizar ou programar suas ações frequentes; o sistema precisa ser fácil para usuários leigos, mas flexível para se tornar ágil a usuários avançados.
8	Projeto estético e minimalista: Os diálogos do sistema precisam ser simples e claros; não devem conter informações irrelevantes ou desnecessárias.
9	Suporte aos usuários no reconhecimento, diagnóstico e recuperação de erros: As mensagens de erros devem ter redação simples e clara, indicar o problema e sugerir soluções.
10	Informações de ajuda e documentação: A documentação do sistema deve ser facilmente acessível, com foco nas tarefas comuns do usuário e não pode ser muita longa.

Quadro 1 – As Dez Heurísticas de Usabilidade de Nielsen

Fonte: Adaptado de Dias (2007).

Além de identificar os problemas, os avaliadores também classificam os problemas de usabilidade encontrados de acordo com o seu grau de severidade. O objetivo é identificar aqueles que são mais sérios e deverão ser corrigidos prioritariamente. Neste trabalho, a classificação foi realizada utilizando uma escala com valores variando entre números de 1 (um) a 4 (quatro), sendo que o menor número indica um problema cosmético e o maior indica um problema catastrófico conforme a Tabela 28.

Tabela 28 - Níveis de Severidade dos Problemas de Usabilidade

Descrição	Nível de Severidade
Problema Cosmético: não há urgência na correção do problema, pois não há interferência no funcionamento do sistema.	1
Problema Simples: a prioridade de correção é baixa, pois o usuário pode facilmente contornar o problema.	2
Problema Grande: a prioridade de correção é alta, pois o usuário terá dificuldade de concluir a tarefa.	3
Problema Catastrófico: o problema impede a conclusão da tarefa e correção deve ser realizada.	4

Fonte: Nielsen (1994).

Cada avaliador associa todos os problemas de usabilidade encontrados às heurísticas violadas e apresentam os resultados em um relatório. Os relatórios individuais são posteriormente consolidados em um único relatório de problemas de usabilidade.

2.3 AVALIAÇÃO AUTOMÁTICA DE USABILIDADE

Diversas técnicas de avaliação de usabilidade têm sido propostas para identificar problemas de usabilidade diferentes e em estágios distintos do desenvolvimento de *software*. Cada técnica tem seus próprios requisitos e, geralmente, técnicas diferentes descobrem problemas de usabilidade diferentes (IVORY; HEARST, 2001; FERNANDEZ et al., 2011; WALJI, 2013).

Estudos apontam que a mesma técnica aplicada numa mesma interface por avaliadores diferentes resultam em problemas de usabilidade distintos. Isso indica que não há uma forma sistemática de descobrir problemas de usabilidade. Essa característica, aliada ao fato de que cada técnica descobre apenas alguns problemas, é o que leva à recomendação para o uso de várias técnicas diferentes (MOLICH et al., 1999; NIELSEN, 1994; FERNANDEZ et al., 2011).

Uma das alternativas para aprimorar essa falta de sistemática e ampliar a faixa de cobertura dos problemas descobertos é automatizar alguns aspectos ou etapas da avaliação de usabilidade, por exemplo, a forma de captura dos dados, a análise da informação coletada e a sugestão de melhorias (IVORY; HEARST, 2001). A avaliação automática de usabilidade apresenta algumas vantagens em relação aos métodos tradicionais de avaliação (NIELSEN, 1994):

- Redução do custo da avaliação. Métodos que automatizam a captura, a análise e a etapa de sugestão de melhorias podem diminuir o tempo gasto na avaliação e, conseqüentemente, o custo.
- Aumento da consistência dos erros descobertos. Em certos casos, podem ser implementadas ferramentas para descobrir desvios nas tarefas comparando-as com modelos definidos como ideais. Isso também permite detectar os padrões de comportamento do usuário que podem sugerir possíveis erros.
- Redução da necessidade de avaliadores experientes. Automatizar algumas partes do processo de avaliação como a análise e a sugestão de melhorias pode ser uma forma de ajudar os avaliadores que não têm muita experiência.
- Aumento da faixa de cobertura dos erros descobertos. Fatores como tempo, custo e recursos restritos não permitem que sejam avaliados todos os aspectos de uma interface. Ferramentas que identificam possíveis padrões de comportamento dos usuários podem aumentar a quantidade de características de usabilidade avaliadas.
- Comparação entre interfaces. Algumas abordagens para avaliação automática permitem realizar comparações entre interfaces distintas, aprimorando o processo de *design* de interface.
- Avaliação da usabilidade durante a fase de *design* de interface. A maior parte das técnicas de avaliações tradicionais permite a avaliação de usabilidade somente após a fase de *design* ou prototipação do projeto, o que aumenta o custo do projeto.

Apesar de ser utilizado o termo “avaliação automática”, a maior parte das técnicas ou ferramentas para avaliação automática de usabilidade requer um esforço humano expressivo porque dependem de testes tradicionais ou exigem intensa interação do avaliador (BALBO et

al., 2005). Nesse sentido, uma classificação da avaliação de usabilidade pode ser dividida em quatro abordagens:

- Não automática: realizada por avaliadores especialistas.
- Captura automática: uso de *software* para facilitar o registro de informação relevante sobre o usuário e o sistema.
- Análise automática: métodos que identificam os problemas de usabilidade automaticamente.
- Análise crítica: métodos que, além de identificar os problemas automaticamente, também sugerem melhorias.

Predominantemente, a automação dos testes de usabilidade tem sido utilizada em duas etapas: na captura dos dados e na análise desses dados de acordo com alguma métrica ou modelo. Fernandez et al. (2011) apresentaram um estudo sistemático dos métodos de avaliação relatados em 206 trabalhos sobre avaliação de sistemas web. Os resultados mostraram que 58% das avaliações foram realizadas através de testes com usuário, mas apenas 31% foram realizadas através de avaliação automática. Os métodos não foram aplicados com exclusividade, pois em um mesmo trabalho podem ter sido aplicados um ou mais métodos. Os autores dividem os métodos de avaliação de usabilidade de sistemas web em cinco categorias:

1. Testes com usuários: envolvem o avaliador observando os usuários na interação com o sistema. Exemplos: análise de *log* e protocolo *think-aloud*.
2. Métodos de inspeção: envolvem um avaliador experiente usando um conjunto de critérios para identificar potenciais problemas de usabilidade. Exemplos: avaliação heurística e regras de recomendação.
3. Investigação: envolvem a aplicação de perguntas subjetivas sobre a interação do usuário com o sistema. Exemplos: entrevistas e questionários.
4. Modelos analíticos: envolvem alguma abordagem que permite ao avaliador prever a usabilidade de um sistema aplicando modelos. Exemplos: família GOMS e testes cognitivos.
5. Simulação: envolvem a simulação da interação do usuário com o sistema através da utilização de algoritmos de simulação. Exemplo: modelo Petri *Net*.

Dos 206 estudos analisados, 69% das avaliações foram realizadas por métodos manuais e 31% utilizaram alguma forma de avaliação automática. São considerados métodos automáticos: análise de *log*, simuladores, algoritmos para análise de código, dentre outros. Por outro lado, exemplos de métodos manuais são: entrevistas, questionários e protocolo *think-aloud*. Na avaliação automática de usabilidade, os avaliadores apenas precisam interpretar os resultados depois que as atividades principais do método foram realizadas automaticamente. A interpretação realizada pelo avaliador pode ser utilizada na tomada de decisão em relação ao sistema, por exemplo, se o sistema está em conformidade com os requisitos ou existe alguma necessidade de reprojeto.

2.4 PADRÕES PARA AVALIAÇÃO DA QUALIDADE DE SOFTWARE

A série de padrões ISO/IEC 25000, também conhecida como série SQuaRE (*System and Software Quality Requirements and Evaluation*), teve o objetivo de criar um *framework* para a avaliação da qualidade de produto de *software*. Esse *framework* é o resultado da evolução de vários outros padrões, especificamente da ISO/IEC 9126 e da ISO/IEC 14598, que definem um modelo de qualidade de produto de *software* e um processo de avaliação da qualidade, respectivamente.

A série SQuaRE é composta por cinco divisões, conforme apresentado na Figura 2.

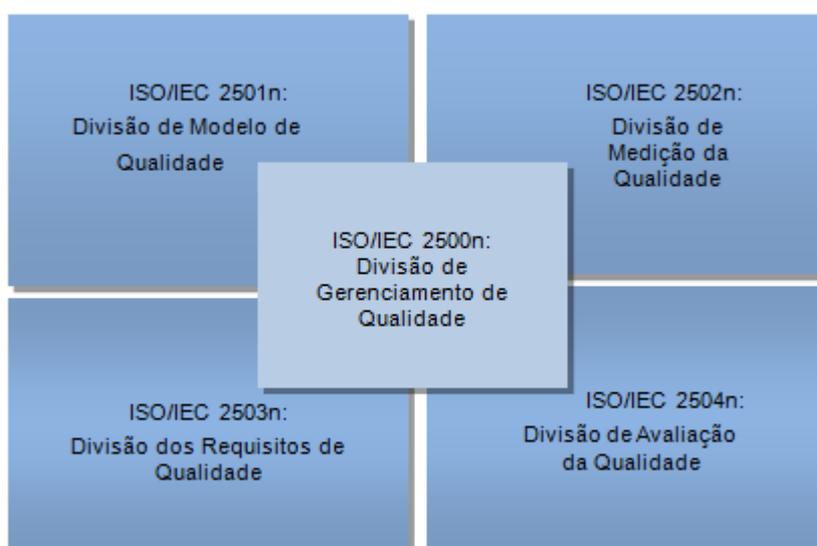


Figura 2 – Divisões da Série SQuaRE
Fonte: Adaptado de International... (2014).

Neste trabalho foram utilizadas as normas: ISO/IEC 25010, que apresenta modelos de qualidade de *software*, qualidade em uso e dados; ISO/IEC 2502n, que aborda métricas de *software*, com definições matemáticas de métricas de qualidade; e ISO/IEC 25040, que especifica requisitos, recomendações e diretrizes para avaliação de produto de *software*.

A definição de um modelo de qualidade é tratada na norma ISO/IEC 25010, que divide o modelo em duas partes: modelo de qualidade de produto de *software* e modelo de qualidade em uso.

O modelo de qualidade de produto de *software* é composto por oito características, com cada característica dividida em subcaracterísticas mensuradas interna ou externamente. Essas características ou subcaracterísticas são medidas externamente quando o *software* faz parte de um sistema de computador ou internamente quando estão relacionadas a atributos do *software* e ao comportamento do sistema. As características do modelo de qualidade de produto de *software* estão representadas na Figura 3, conforme a definição da ISO/IEC 25010.

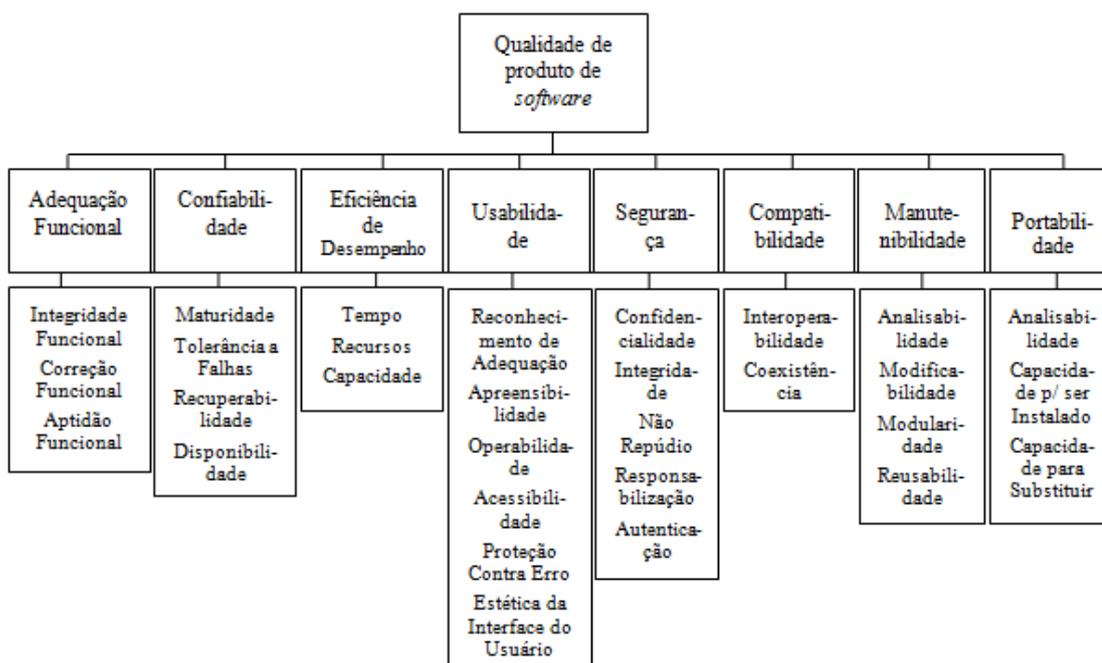


Figura 3 – Características de Qualidade de Produto de *Software*

Fonte: Adaptado de International...(2011).

A qualidade em uso é definida por outras três características principais, representadas na Figura 4, mensuradas no contexto real ou na simulação de uso do sistema: eficácia em uso, eficiência em uso, satisfação em uso e usabilidade em uso. A definição de um modelo de qualidade fornece a base para especificação dos requisitos de qualidade e para a avaliação da

qualidade, que podem ser utilizados sobre diferentes perspectivas: desenvolvedores, clientes e demais interessados em especificar e avaliar a qualidade de um *software*.

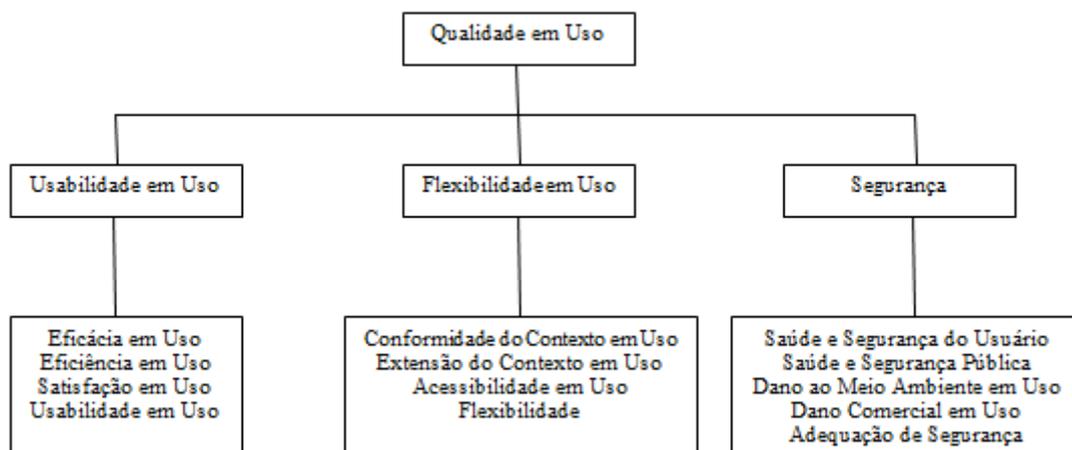


Figura 4 – Características de Qualidade em Uso

Fonte: Adaptado de International...(2011).

Em relação à avaliação da qualidade, a ISO/IEC 25040 define requisitos, recomendações e diretrizes para a avaliação do produto de *software*, especificando um modelo de referência do processo, com informações genéricas para a avaliação da qualidade. No modelo de referência do processo de avaliação, a norma prevê cinco etapas que devem ser seguidas para avaliar um *software*. Para cada uma dessas etapas existem outras tarefas, cada qual gerando saídas (documentação) sobre o processo:

1. Estabelecer os requisitos da avaliação:
 - a. Estabelecer o propósito da avaliação;
 - b. Obter os requisitos de qualidade do produto de *software*;
 - c. Identificar as partes do produto a serem incluídas na avaliação;
 - d. Definir o rigor da avaliação.

Saída: Requisitos da avaliação;

2. Especificar a avaliação:
 - a. Selecionar as métricas;
 - b. Definir os critérios de decisão para as métricas;
 - c. Definir os critérios de decisão para avaliação.

Saída: Especificação da avaliação.

3. Projetar a avaliação:
 - a. Projetar as atividades de avaliação;

Saída: Projeto da avaliação.

4. Executar a avaliação:

- a. Obter os valores das métricas;
- b. Aplicar critérios de decisão para as métricas;
- c. Aplicar critérios de decisão para avaliação.

Saída: Resultados da avaliação.

5. Concluir a avaliação:

- a. Revisar os resultados;
- b. Dispor os dados.

Saída: Conclusão da avaliação.

Para as tarefas de avaliação, o Anexo D (Informativo) da Norma ISO/IEC 25040 ainda sugere técnicas de avaliação indicadas para cada característica de qualidade. Em relação à usabilidade, a norma define três técnicas: inspeção da documentação e da interface do usuário, verificação da conformidade em relação a padrões de interface e testes com usuários.

O modelo de referência do processo de avaliação produto de *software* está representado na Figura 5.

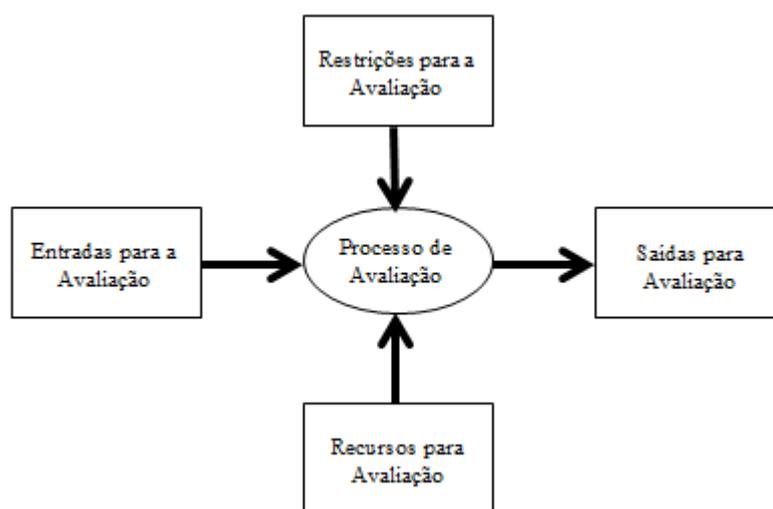


Figura 5 – Modelo de Referência/ Avaliação de Qualidade de Produto de Software
 Fonte: Adaptado de International...(2011).

Com o processo de avaliação definido de forma sistemática e pautado em um modelo de qualidade previamente estruturado, é possível avaliar a qualidade de um produto de *software* utilizando definições matemáticas. As normas ISO/IEC 2502n descrevem um modelo de referência para mensurar a qualidade de um *software*, apresentam definições

matemáticas para cálculo dessas métricas e oferecem um guia prático para a aplicação. O modelo de referência para mensurar a qualidade está esquematizado na Figura 6.

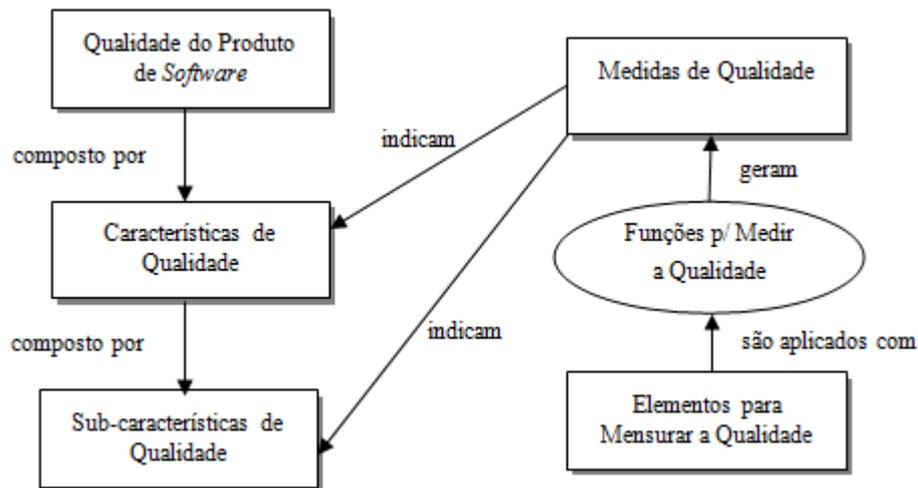


Figura 6 – Modelo de Referência para Mensurar a Qualidade do Produto de Software
 Fonte: Adaptado de International...(2007).

Um método para mensurar a qualidade é uma sequência lógica de operações usadas para quantificar propriedades e relacioná-las a uma escala especificada (INTERNATIONAL..., 2012b). As características e subcaracterísticas podem ser quantificadas aplicando funções para mensurar a qualidade. Uma função é um algoritmo usado para combinar elementos de qualidade. O resultado da execução desses algoritmos é chamado de métrica de qualidade de *software*. Dessa forma, métricas de qualidade se tornam quantificações de características e subcaracterísticas de qualidade.

Em relação à característica de usabilidade, as métricas de qualidade de um *software* podem ser usadas para indicar o quanto esse *software* é amigável, fácil de operacionalizar, entender e aprender. Além disso, também é possível mensurar essa característica enquanto os usuários testam uma determinada função. Nesse caso, os resultados serão influenciados pela capacidade dos usuários e pelo contexto de uso. No entanto, esse fato não invalida os resultados, desde que a avaliação seja executada sobre condições especificadas para um conjunto de usuários que possa representar ou identificar um grupo. Para resultados consideráveis, pelo menos oito usuários são necessários, embora informação útil possa ser obtida de grupos menores (INTERNATIONAL..., 2012c).

2.5 CAPTURA DA INTERAÇÃO DO USUÁRIO

A captura da interação do usuário é muito utilizada para obter dados que serão utilizados para avaliação de usabilidade e estudos do comportamento do usuário (PREECE et al., 2012; TURNBULL, 2006; JANSEN, 2006; KELLAR; WATTERS, 2005). Esses dados também podem ser usados para justificar a realização de manutenção ou alterações nos sistemas.

Sistemas web que utilizam a tecnologia RIA são projetados com uma ou mais páginas, compostas por componentes individuais, que podem ser independentemente atualizados, excluídos ou adicionados em tempo de execução (AMALFITANO et al., 2010). O emprego de tecnologia RIA faz com que os sistemas web sejam similares aos sistemas implementados para uso local (*desktop*), considerando os recursos disponíveis para interatividade do usuário, trazendo ganho na característica de usabilidades desses sistemas (AMALFITANO et al., 2010).

Neste trabalho foi utilizada a *Web Application Usage Tracking Tool* (WAUTT) (RIVOLLI et al., 2014), uma ferramenta para capturar informações detalhadas sobre a interação do usuário com uma RIA. A WAUTT é utilizada para capturar a interação do usuário com sistemas web em um computador cliente e registrar todos os eventos coletados em um servidor para serem usados para avaliação de usabilidade. A captura dos dados ocorre com a inserção de um código *JavaScript* nas páginas da RIA. Além de capturar um grande número de detalhes, a ferramenta permite ao avaliador selecionar eventos específicos de acordo o objetivo da avaliação.

Os dados de interação capturados pela WAUTT são chamados de eventos. Considera-se um evento qualquer interação do usuário com um elemento específico da interface, por exemplo: o clique e o movimento do *mouse* (considerando a coordenada em relação ao canto superior esquerdo da tela), a rolagem da página (considerando a posição da barra de rolagem a partir do topo da tela), a mudança de foco sempre que um elemento receber ou perder o foco, o carregamento completo da página, o redimensionamento da janela do navegador, dentre outros.

A ferramenta foi projetada para capturar e armazenar os dados em um banco de dados relacional. As informações armazenadas no banco de dados podem ser acessadas por meio da interface administrativa. Cada interação do usuário pode ser associada a um dos seguintes grupos de informação:

- Dados da sessão: inclui o identificador único da sessão, sistema operacional, navegador e endereço IP da máquina. A sessão refere-se à sequência de ações do mesmo usuário durante o uso da aplicação e é criada quando o usuário abre o navegador e carrega o aplicação. A sessão encerra-se quando o navegador é fechado.
- Dados da navegação: inclui o título da página, endereço, tamanho e tipo de requisição (*get* ou *post*). Os dados de navegação envolvem a sequência de páginas carregadas durante o uso do sistema. Enquanto em alguns sistemas os usuários interagem com uma única página sem atualizá-la completamente, em outros sistemas várias páginas são carregadas.
- Dados de evento: inclui os detalhes específicos de cada ação executada pelo usuário na interface e variam de acordo com o tipo de evento.
- Dados de atributos (DOM): inclui o nome do elemento da página e seus atributos, por exemplo, *id*, *class*, *title* e atributos específicos, por exemplo, *value* e *name* para o elemento tipo *input*.
- Dados da tarefa: essa informação refere-se à tarefa ativa, caso a ferramenta esteja monitorando várias tarefas para permitir a realização da análise dos dados capturados por tarefa.

A Figura 7 mostra dados capturados pela WAUTT e acessados via interface administrativa.



Figura 7 – Interface administrativa para acesso aos dados capturados pela WAUTT
Fonte: Rivolli et al. (2014).

A captura de dados de eventos de forma automática, como a realizada pela WAUTT, é um ponto importante para a avaliação de usabilidade descrita neste trabalho. No entanto, é necessário executar outras atividades para analisar e interpretar esses dados, obtendo, dessa forma, conhecimento útil para mensurar a qualidade de um *software*. A mineração dos dados capturados, descrita na Seção 2.6, é uma das atividades necessária para iniciar a análise dos dados, com a finalidade de auxiliar os avaliadores a identificar os dados considerados significativos, reduzindo o volume de dados e focalizando a tomada de decisão.

2.6 MINERAÇÃO DE DADOS

A Mineração de Dados (*Data Mining*) em documentos web, também conhecida como *Web Mining*, refere-se à aplicação de técnicas de mineração para automatizar a descoberta e extração de conhecimento útil em grandes bases de dados web (HAN; KAMBER, 2011). Em geral, a mineração de dados na web abrange três campos de pesquisa: mineração de conteúdo, mineração de estrutura e mineração de uso.

A mineração de conteúdo refere-se à descoberta de conhecimento contido nos documentos web, que podem ser de formatos variados (texto, imagem, áudio e vídeo). A mineração de estrutura descobre conhecimento útil dos *hiperlinks*, que representam a estrutura da web. A mineração de uso trabalha com as informações de como os usuários interagem com o sistema web. Esse campo de pesquisa também é conhecido como *Web Log Mining* e refere-se à descoberta de padrões de comportamento dos usuários. Enquanto a mineração de conteúdo e a mineração de estrutura utilizam os dados primários da web, a mineração de uso utiliza dados secundários provenientes de arquivos de *log*, perfis de usuário, pasta favoritos (*bookmarks*), ou quaisquer outros dados de interação (SHARMA et al., 2011).

Os métodos aplicáveis à mineração de dados na web são definidos de acordo com os dados que serão selecionados e os resultados a serem obtidos. Dentre os métodos que podem ser utilizados, destacam-se: análise estatística, associação, agrupamento (*clustering*), classificação, predição e regressão (ASHRAF; ASHRAF, 2014).

Neste trabalho, o método utilizado foi o *clustering*, que é o estudo formal de algoritmos e métodos para agrupar objetos com características semelhantes (MIRKIN, 2012). Diferentemente do método de classificação, que trabalha com objetos com características conhecidas e sobre as quais é possível estabelecer uma descrição ou rótulo, no *clustering* não se sabe inicialmente quais serão essas classes. A finalidade do *clustering* é encontrar classes

que são diferentes umas das outras, mas que possuam membros similares entre si, segundo alguma medida de similaridade.

O *clustering* é uma técnica de mineração de dados conhecida como aprendizado não supervisionado. Enquanto no aprendizado supervisionado é possível conhecer previamente os objetos (atributos/entrada) e suas classes (saída), no aprendizado não supervisionado, primeiro é necessário descobrir padrões e similaridades entre os objetos para depois definir as classes.

Definições formais de um agrupamento (*cluster*) são (EVERITT et al., 2011):

1. Objetos de um mesmo *cluster* são semelhantes entre si e diferentes dos objetos dos outros agrupamentos;
2. Um *cluster* é um conjunto de ponto no espaço tal que a distância entre os pontos em um mesmo conjunto é menor que a distância entre pontos de diferentes aglomerados;
3. Os *clusters* podem ser descritos como regiões conexas de um espaço multidimensional que contém uma grande densidade relativa de pontos. As regiões estão separadas umas das outras por regiões de baixa densidade relativa de pontos.

De forma geral, o processo de *clustering* é composto por quatro etapas (JAIN; DUBES, 1988):

1. Aquisição dos dados (seleção das amostras, seleção das características de cada amostra, pré-processamento dos dados);
2. Definição da medida de similaridade (por exemplo, distância euclidiana, correlação de Pearson e distância de Manhattan);
3. Aplicação do algoritmo de *clustering*;
4. Análise e interpretação dos *clusters*.

Algoritmos de *clustering* podem ser aplicados nos três campos de pesquisa da mineração de dados na web (mineração de conteúdo, de estrutura e de uso). O foco deste trabalho foi a mineração de uso. Apesar dos dados constantes dos arquivos de *log* de eventos serem conhecidos, não é possível saber exatamente o conhecimento útil que será extraído a partir da sua análise.

Neste trabalho o algoritmo escolhido para executar o *clustering* foi o algoritmo X-Means, uma extensão do K-Means (MACQUEEN, 1997). Embora o K-Means tenha sido proposto a mais de 50 anos, ainda é um dos algoritmos de *clustering* mais utilizados. A facilidade de implementação, a simplicidade, a eficiência e a diversidade de campos de aplicação são algumas das razões da sua popularidade (JAIN, 2010).

O objetivo do K-Means é escolher k objetos (aleatoriamente ou seguindo algum critério) que serão definidos como centroides (centro do *cluster*). Todos os demais objetos são associados ao centroide mais próximo. A cada iteração do algoritmo, os centroides são recalculados dentre os objetos de seu próprio *cluster* e os objetos são realocados para o centroide mais próximo. Esse passo é repetido até que o ponto de convergência seja satisfatório, ou seja, até que não haja mais alteração na definição dos centroides.

No entanto, o K-Means tem uma limitação relacionada à definição do número de *clusters* (k): o avaliador é quem precisa definir o número k , o que força o algoritmo a trabalhar com um número fixo de *clusters*. O fato de fixar esse número pode resultar em um ponto de convergência insatisfatório. Por esse motivo, foi escolhido o X-Means, que define dinamicamente o número de *clusters*, na busca por um ponto de convergência considerado ótimo (PELLEG, 2000).

3 TRABALHOS RELACIONADOS

Nas próximas seções são apresentados os trabalhos pesquisados na literatura que estão relacionados à avaliação de usabilidade automática realizada através do método de análise de *log*. O capítulo está dividido em duas seções. A Seção 3.1 apresenta um resumo das principais características dos trabalhos. A Seção 3.2 descreve a análise e a comparação dos trabalhos relacionados quanto a essas características.

3.1 ANÁLISE DE LOG PARA AVALIAÇÃO DE USABILIDADE

O Quadro 2 mostra um resumo dos trabalhos que utilizam a análise de *log* como método para avaliar a usabilidade de sistemas web. Os trabalhos são apresentados em ordem cronológica.

Referência	Descrição
WebRem Usine (PAGANELLI; PATERNÒ, 2002)	A ferramenta apresentada no trabalho permite a análise de <i>logs</i> de interação do usuário utilizando um modelo de tarefas definido para uma aplicação específica. O método é composto por 3 fases: a) preparação, com a definição do modelo de tarefas; b) análise de <i>log</i> automática, fornecendo as tarefas executadas, o número de erros e tempo de conclusão; c) avaliação, com a análise humana para visualizar os problemas encontrados e sugerir melhorias. A ferramenta apresenta os resultados de forma gráfica, indicando quantas vezes e quais tarefas foram executadas com sucesso e com erro, a lista das tarefas que nunca foram executadas corretamente e o padrão de comportamento dos usuários em tarefas realizadas com sucesso. Os autores finalizam o artigo propondo melhorias na ferramenta para minimizar os esforços na fase de preparação do método.
AWUSA (TIEDTKE et al., 2002)	Os autores apresentam uma ferramenta que utiliza diferentes técnicas de captura, e mineração de dados para fazer a avaliação automática de usabilidade em um sistema web. Não há a presença de um avaliador para fazer análise humana dos resultados nem sugestão de melhorias da interface. A abordagem da ferramenta prevê: a) visualização dos caminhos percorridos pelo usuário; b) visualização das tarefas não concluídas; c) visualização das interfaces ou situações com pouca usabilidade; d) visualização de novas tarefas que usuários têm a intenção de executar; e) agrupamento dos usuários de acordo com suas tarefas/objetivos. Com essas informações e através da utilização de algoritmos de agrupamento e regras de associação, os problemas de usabilidade encontrados são classificados em cinco categorias: interação, informação, arquitetura, design e técnico. Os autores concluem que a automação é uma vantagem comparada às abordagens de avaliação de usabilidade tradicionais. No entanto, os métodos tradicionais também têm vantagens significativas e o ideal seria criar um sistema que combinasse a avaliação clássica com uma abordagem automatizada.

Referência	Descrição
WAUTER (BALBO et al., 2005)	O estudo apresenta a WAUTER, uma ferramenta com código aberto que faz a captura da interação do usuário e compara-os com um modelo de tarefas previamente definido. O projeto é composto por quatro ambientes com funções distintas: 1) captura dos dados; 2) definição do modelo de tarefas; 3) produção automática de anotações nos arquivos de <i>log</i> e no modelo de tarefas; 4) criação de um modelo visual para apoio à avaliação de usabilidade realizada por especialistas. A ferramenta não faz análise crítica dos dados, mas fornece informações para que o avaliador a faça. Os autores finalizam o estudo propondo uma extensão da WAUTER para fornecimento de informação estatística sobre os dados coletados para facilitar a análise do avaliador.
Ergo-Monitor (SCHWERZ et al., 2007).	A ferramenta Ergo-Monitor é um protótipo de sistema de monitoramento que utiliza as informações armazenadas nos arquivos de <i>log</i> de servidor para realizar uma comparação com um modelo de tarefas previamente definido. O resultado da comparação das informações é apresentado através de laudos quantitativos que relacionam taxas e métricas de usabilidade. Esses valores devem, posteriormente, ser confrontados pelo avaliador com valores que indiquem características de outros sistemas web considerados de boa qualidade.
Usaproxy (ATTERER; SCHIMIDT, 2007)	A abordagem apresentada pela Usaproxy utiliza modelo de tarefas para comparação dos dados de interação do usuário. Os dados coletados apresentam informações detalhadas sobre a interação com o usuário, capturados através de um servidor <i>proxy</i> . No entanto, não detecta atributos HTML ou CSS.
WebHint (VARGAS et al., 2010)	WebHint faz a análise de usabilidade através da captura de dados de interação do usuário, aplicação de um algoritmo de <i>Split log</i> e técnicas de agrupamento. Diferentemente das abordagens que apenas comparam sequência de passos, WebHint utiliza padrões de comportamentos mais comuns para compará-los com o modelo de tarefas que, segundo os autores, fornece uma melhor compreensão do comportamento dos usuários.
Google Analytics (PLAZA, 2011)	No estudo apresentado, o autor utiliza os dados capturados pela ferramenta Google Analytics para mensurar o desempenho de websites. Os dados gerados foram exportados e analisados por <i>software</i> estatístico. As informações referem-se às páginas visitadas, tempo de permanência e caminho percorrido. Não há informação sobre eventos de interação com o usuário. A ideia do autor é aprimorar a forma de interpretação dos dados para centralizar os estudos na geração de informação referente à característica de eficácia dos sistemas web.
Usabilics (VASCONCELOS, BALDOCHI, 2011)	A ferramenta Usabilics é baseada na comparação de sequência entre as tarefas realizadas pelo usuário e um modelo de tarefas. O destaque dado pelos autores refere-se à facilidade de definir o modelo de tarefas de forma intuitiva – o avaliador seleciona a tarefa a ser definida e grava os passos realizados apenas navegando pelo sistema. A abordagem faz o cálculo de similaridade entre as tarefas executadas e esperadas através do UsabilityIndex. Conhecendo os problemas e o resultado dos índices, o avaliador pode propor melhorias na interface, recalcular os índices e analisar se as alterações foram realmente válidas.
Web Usability Probe (CARTA et al., 2011)	A abordagem proposta pela ferramenta Web Usability Probe utiliza um servidor <i>proxy</i> para capturar a interação do usuário com o sistema web. O avaliador, ao acessar esse servidor, além de utilizar os dados para conseguir informações sobre a usabilidade, pode configurar a forma como esses dados serão apresentados nos relatórios com o objetivo de facilitar a leitura da informação. Um dos objetivos da ferramenta é facilitar a comparação da interação de vários usuários para melhor identificar comportamentos divergentes daquele considerado como ótimo. O foco do trabalho está na captura de dados detalhados e na possibilidade de visualizar esses dados de forma gráfica, o que facilita o trabalho do avaliador na busca de problemas de usabilidade.

Referência	Descrição
Web Log Mining Tool (DUAN et al., 2012)	A ferramenta trabalha com conceitos de <i>Web Mining</i> , utilizando o processo de descoberta de conhecimento em arquivos de <i>log</i> de servidor. Os autores afirmam que a abordagem proposta é capaz de identificar padrões de comportamento dos usuários, frequência e conteúdos acessados. A análise de <i>log</i> é realizada utilizando algoritmos de regras de associação (Apriori) e algoritmos <i>sequential pattern mining</i> . O resultado é a construção de grafos para comparação com um modelo de tarefas.
Welfit (SANTANA, BARANAUSKAS; 2015)	O trabalho descreve uma técnica para processamento de arquivos de log e construção de gráficos com informações estatísticas sobre o padrão de comportamento dos usuários. Através dos gráficos, avaliadores especialistas podem encontrar problemas de acessibilidade e usabilidade. Utiliza o algoritmo SAM (Sequence Alignment Method) para verificar a quantidade de operações necessárias para concluir uma tarefa. As vantagens da abordagem apresentada são a utilização de dados detalhados e a não utilização de modelo de tarefas. No entanto, para avaliar os possíveis problemas de usabilidade ou acessibilidade, o avaliador deve conhecer os eventos de interface do usuário (UI) e os elementos da página web. Os autores sugerem a criação de mecanismos que ajustem automaticamente os problemas apontados pela ferramenta e eliminação dos padrões de comportamento que representem respostas erradas do usuário (falsos positivos).

Quadro 2 – Trabalhos Relacionados

Fonte: Autoria própria.

3.2 ANÁLISE E COMPARAÇÃO

Da análise dos trabalhos relacionados, algumas características comuns podem ser destacadas: utilização de um modelo para fazer comparação entre tarefas; captura das informações da interação do usuário com o sistema; e análise e interpretação dos dados. A seguir, as ferramentas são analisadas segundo essas características. Ao final, um quadro comparativo é apresentado.

As ferramentas AWUSA, Ergo-Monitor, Usaproxy, Usabilics, WAUTER, WebRemUsine, WebHint, Web Usability Probe e Web Log Mining Tool utilizam um modelo de tarefas para verificar a similaridade entre a sequência de passos realizados pelo usuário durante a execução de uma tarefa específica e a sequência definida pelo avaliador, considerada como modelo. A diferença entre as ferramentas está no modo de definir o modelo. A ferramenta Usabilics é a única que define o modelo de tarefas de forma interativa, apenas gravando a navegação do avaliador no sistema web. As outras ferramentas realizam essa etapa utilizando notações e análise da topologia do sistema web para representar uma tarefa de forma gráfica.

As ferramentas Welfit e Google Analytics, diferentemente da abordagem anterior, não definem previamente um modelo de tarefas. A Welfit utiliza a comparação entre sequências através da aplicação do algoritmo SAM (*Sequence Alignment Method*). O SAM implementa uma técnica de classificação que calcula a distância entre a quantidade de operadores necessários para concluir uma tarefa e o caminho realizado pelo usuário. Quanto menor a distância entre as duas sequências, maior a similaridade e menor o problema de usabilidade. No entanto, a ferramenta não considera o caminho entre as páginas de um sistema web, mas apenas identifica os padrões nas ações do usuário em uma única página.

Em relação à captura da interação do usuário com o sistema, as ferramentas que registram os eventos relacionados à interação do usuário com o sistema são: Usabilics, UsaProxy, WAUTER, Web Usability Probe, WebHint, Welfit. As duas últimas não capturam atributos CSS e atributos HTML específicos dos elementos da página.

Considerando a forma de análise e interpretação dos dados, tanto a ferramenta Web Log Mining Tool quanto a WELFIT utilizam algoritmos de análise de grafos para a representação da navegação do usuário pelo sistema de forma gráfica. Em ambos os casos, o avaliador deve analisar os grafos gerados na busca de problemas que possam indicar falta de usabilidade.

Nas ferramentas AWUSA e WebHint, técnicas de agrupamento são aplicadas nos arquivos de *log*, com o objetivo de agrupar os padrões de comportamento de usuários com características semelhantes para facilitar a análise pelos avaliadores. Na WebHint, o sistema utiliza um algoritmo denominado *Split Log* para realizar a divisão das tarefas de cada usuário em um arquivo separado e, em seguida, aplica o algoritmo de agrupamento para encontrar padrões comuns de comportamento e compará-los com heurísticas, detectando possíveis problemas de usabilidade. A diferença entre WebHint e AWUSA é que a última utiliza algoritmos de regras de associação após o agrupamento.

Usabilics faz a análise dos dados capturados através do cálculo do Índice de Usabilidade (*UsabilityIndex*). O cálculo leva em consideração a sequência de eventos realizada pelo usuário, a completude da tarefa (quando houve êxito na execução), a quantidade eventos por tarefa e o total de tarefas executadas. O resultado desse índice permite ao avaliador uma visualização rápida e resumida sobre a usabilidade da interface, além de fornecer informação para comparar os valores dos índices antes e depois de alguma readaptação da interface.

As ferramentas Google Analytics e Web Usability Probe utilizam *software* estatístico e relatórios descritivos para apresentar dados resumidos sobre a interação do usuário com o

sistema. A primeira utiliza uma abordagem com foco no tempo de permanência no sistema e páginas navegadas. A Web Usability Probe permite ao avaliador configurar diversas formas de apresentação de relatórios com dados relacionados à interação do usuário.

Não há informação da forma como os dados são processados com a Usaproxy. O trabalho apenas apresenta a forma capturar a interação do usuário através da configuração do navegador do cliente ou de um *proxy*. Com a ferramenta WebRemUsine, além da captura da interação também é apresentado o modo de processamento da informação, que é realizado por meio de uma tabela específica de mapeamento entre as entradas de *log* e o modelo de tarefas.

O Quadro 3 apresenta uma síntese comparativa com as principais características de cada uma das ferramentas. A coluna “Modelo de Tarefas” indica se a ferramenta utiliza um modelo para comparar a sequência de passos realizados pelo usuário com a sequência considerada como ideal e definida pelo avaliador. A coluna “Captura de Eventos” indica se é realizada a captura dos eventos de interação do usuário com o sistema. A coluna “Análise Automática” refere-se à forma de interpretação dos dados analisados. A análise é considerada “Parcial” quando existe a necessidade do avaliador participar da elaboração dos resultados e “Total” quando o resultado da avaliação de usabilidade apresenta dados conclusivos, sem a influência humana no processo de interpretação dos resultados da análise de *log*. As informações referentes à coluna “Processamento” indicam quais técnicas ou ferramentas são utilizadas para interpretar os *logs*. A coluna “Características” indica algumas características das ferramentas, consideradas relevantes, que auxiliaram para justificar a importância da abordagem apresentada neste trabalho.

Ferramenta	Modelo Tarefas	Captura de Eventos	Análise Automática	Processamento	Características
WebRem Usine	Sim	Não	Parcial	Comparação de sequências	Não é transparente ao usuário
AWUSA	Sim	Não	Total	Agrupamento e Regras de Associação	Dificulta a associação de caminhos alternativos para execução de uma única tarefa
WAUTER	Sim	Sim	Parcial	Comparação de sequências de forma visual	Não é transparente ao usuário
Ergo-Monitor	Sim	Não	Parcial	Comparação de sequências	A construção do modelo é baseada exclusivamente em URLs
Usaproxy	Sim	Sim	Total	-	O <i>proxy</i> pode limitar a captura de informação importante
Welfit	Não	Sim	Parcial	Algoritmos de Grafos	Não é transparente ao usuário

Ferramenta	Modelo Tarefas	Captura de Eventos	Análise Automática	Processamento	Características
WebHint	Sim	Sim	Parcial	Agrupamento	Utiliza <i>Split Log</i> para comparar modelos
Google Analytics	Não	Não	Parcial	Ferramentas estatísticas	Apresenta relatórios de comportamento do usuário baseados exclusivamente no tempo de navegação e páginas acessadas
Usabolics	Sim	Sim	Parcial	Cálculo do <i>UsabilityIndex</i>	Modelo de tarefas definido de forma intuitiva
Web Usability Probe	Sim	Sim	Parcial	Ferramentas estatísticas	Interface gráfica que permite a elaboração de relatórios com várias estruturas diferentes
Web Log Mining Tool	Sim	Não	Parcial	Algoritmos de Grafos	Os grafos são utilizados para fazer comparação de modelos

Quadro 3 – Características das Ferramentas

Fonte: Autoria própria.

Considerando os estudos apresentados sobre avaliação automática de usabilidade, foi elaborada uma proposta de avaliação de usabilidade de sistemas web apoiada unicamente na interpretação de arquivos de *log* com dados de eventos. As características relevantes da abordagem proposta são:

- Não obrigatoriedade do modelo de tarefas: a busca por informações que podem indicar problemas de usabilidade é realizada exclusivamente pela análise dos arquivos de *log* com dados de eventos. A vantagem de se utilizar um método sem a definição de um modelo de tarefas é a redução da intervenção humana no processo, minimizando os esforços do avaliador e a ocorrência de falhas. De fato, a maior parte dos indicadores de problemas de usabilidade que podem ser mensurados são obtidos unicamente pela análise de *log* (NIELSEN, 1994).
- Utilização de dados de eventos de RIA: a captura dos dados realizada pela WAUTT fornece dados detalhados sobre a interação do usuário com o sistema. Não são registradas somente as páginas visitadas pelo usuário, como em arquivos de *log* de servidor, mas também a interação do usuário com elementos da página. O nível de detalhes obtido por essa captura e a possibilidade de capturar dados de RIA pode fornecer informações adicionais e importantes a respeito da interação do usuário com o sistema.

- Transparência para o usuário: nenhuma das atividades da abordagem exige o conhecimento prévio do usuário. O processo é feito de forma transparente, sem a necessidade de instalação de aplicativos no computador cliente. A participação dos usuários se restringe à navegação pelo sistema avaliado.

4 UM PROCESSO DE AVALIAÇÃO AUTOMÁTICA DE USABILIDADE

Para a definição do processo de avaliação automática de usabilidade apresentada neste trabalho foram consideradas as definições das normas ISO/IEC (Série SQuaRE). Neste capítulo é apresentado um processo de avaliação de qualidade de *software*, representado através da instanciação do *framework* para a avaliação da qualidade de produto de *software* proposto na série SQuaRE.

Este capítulo está dividido em duas seções. A Seção 4.1 apresenta as definições dos conceitos que serão aplicados no decorrer do processo. A Seção 4.2 descreve, de forma genérica, todas as etapas necessárias para execução do processo de avaliação automática de usabilidade proposto.

4.1 DEFINIÇÕES

A avaliação da qualidade de um *software* é um processo sistemático que deve ser muito bem planejado e executado para se alcançar os objetivos definidos (INTERNATIONAL..., 2011). Assim, a norma ISO/IEC 25040 definiu um modelo de referência para a avaliação da qualidade de *software* que delimita o processo de avaliação em entradas, saídas, recursos e restrições. O objetivo é fazer com que o processo de avaliação tenha objetivos claros e critérios definidos.

As entradas utilizadas para o processo de avaliação automática de usabilidade proposto são os arquivos com dados de eventos capturados da interação de usuários com um sistema web e métricas de *software* baseadas nas Normas ISO/IEC 2502n e na literatura (BALBO et al., 2005; NOY et al., 2009; SIOCHIS, 1991; TING et al., 2005; SWALLOW, 1997).

Para a execução do processo são necessários os seguintes recursos: banco de dados relacional para armazenar os dados capturados e usuários que irão utilizar o sistema web. A restrição existente para a execução do processo é que o sistema web usado para a avaliação deve permitir a inclusão de código *JavaScript* em suas páginas. A saída do processo são os indicadores resultantes da aplicação das métricas ao arquivo de *log* com dados de eventos.

4.2 ETAPAS DO PROCESSO

O processo de avaliação de usabilidade proposto neste trabalho foi conduzido seguindo a sequência de cinco etapas definidas pela ISO/IEC 25040: estabelecer os requisitos de avaliação; especificar a avaliação; projetar a avaliação; executar a avaliação; concluir a avaliação. As etapas são descritas nas subseções seguintes.

4.2.1 Estabelecer os requisitos de avaliação

Nesta etapa são previstas as seguintes atividades: estabelecer o propósito da avaliação; obter os requisitos da qualidade de produto de software; identificar as partes do produto a serem incluídas na avaliação; definir o rigor da avaliação.

O objetivo da atividade de estabelecer o propósito da avaliação é tomar decisões sobre a aceitação ou não do sistema, sobre a necessidade de adaptações, comparar ou selecionar sistemas, verificar aspectos positivos e negativos do sistema e decidir quando o sistema deve ser melhorado ou substituído. O propósito deste trabalho é aplicar um processo de avaliação de usabilidade através da análise automática de arquivos de *log* com dados de eventos da interação do usuário para obter indicadores sobre a usabilidade do sistema.

Na atividade de obter os requisitos de qualidade do produto de *software*, devem ser definidas quais as características e subcaracterísticas do sistema serão avaliadas para atender aos requisitos de qualidade. Os requisitos de qualidade para o processo proposto estão limitados às características de qualidade de *software* que podem ser obtidas de forma quantitativa a partir da aplicação de métricas aos arquivos de *log*. Neste trabalho foram utilizadas as seguintes subcaracterísticas de usabilidade: apreensibilidade, operabilidade, eficiência, proteção ao erro do usuário e estética da interface.

Na atividade de identificar as partes do produto que serão incluídas na avaliação, o avaliador deve identificar todas as funções do sistema que serão avaliadas. Para o processo proposto, em geral, as partes são referentes aos elementos interativos da interface.

O rigor da avaliação deve ser definido para garantir confiabilidade na qualidade do produto de *software* levando em consideração o propósito da avaliação. Neste trabalho, dentre as opções previstas na Norma, foi escolhida a técnica de realização de testes com usuários reais.

4.2.2 Especificar a avaliação

Nesta etapa são previstas as seguintes atividades: selecionar as métricas; definir os critérios de decisão das métricas; definir os critérios de decisão para a avaliação.

Na atividade de selecionar as métricas, são definidas apenas as métricas para o processo proposto que podem ser obtidas de maneira quantitativa a partir de um arquivo de *log* com dados de eventos. Dessa maneira, foram consideradas métricas para as subcaracterísticas de apreensibilidade, operabilidade, eficiência, proteção ao erro do usuário e estética da interface, definidas com base em normas internacionais e trabalhos da literatura (INTERNATIONAL..., 2012a; INTERNATIONAL..., 2012b; KRONBAUER; SANTOS, 2011; MORAIS et al., 2013; NOY et al., 2009). O conjunto de métricas definido para o processo (Apêndice A) pode ser adaptado e/ou estendido de acordo com o contexto de uso.

Na atividade de definir os critérios de decisão das métricas, valores limite para cada métrica ou valores considerados como ideais devem ser definidos para indicar a necessidade ou não de ações ou nova investigação com relação à subcaracterística relacionada. As escalas para essas métricas são valores entre 0 (zero) e 1 (um), sendo que os valores mais próximos de 1 (um) sugerem maior conformidade do sistema com a métrica.

Na atividade de definir os critérios de decisão para avaliação, o avaliador deve relacionar os resultados obtidos na etapa anterior (definir os critérios de decisão das métricas) e verificar se todos os indicadores resultantes das métricas estão em conformidade com os valores limite definidos para as subcaracterísticas de usabilidade. Como critério de decisão para avaliação, o sistema pode ser liberado se todas as métricas indicarem conformidade, caso contrário o sistema (ou partes do sistema) deve ser encaminhado para nova investigação através de outros métodos de avaliação de usabilidade.

Os critérios de decisão das métricas e os critérios de decisão para avaliação servem de parâmetro para o avaliador tomar decisões e podem ser alterados de acordo com as peculiaridades de cada sistema.

4.2.3 Projetar a avaliação

Esta etapa consiste em definir e planejar os recursos necessários para aplicação dos testes, considerando as características financeiras, temporais e/ou de recursos humanos. Nesta etapa está prevista a atividade de projetar as atividades de avaliação.

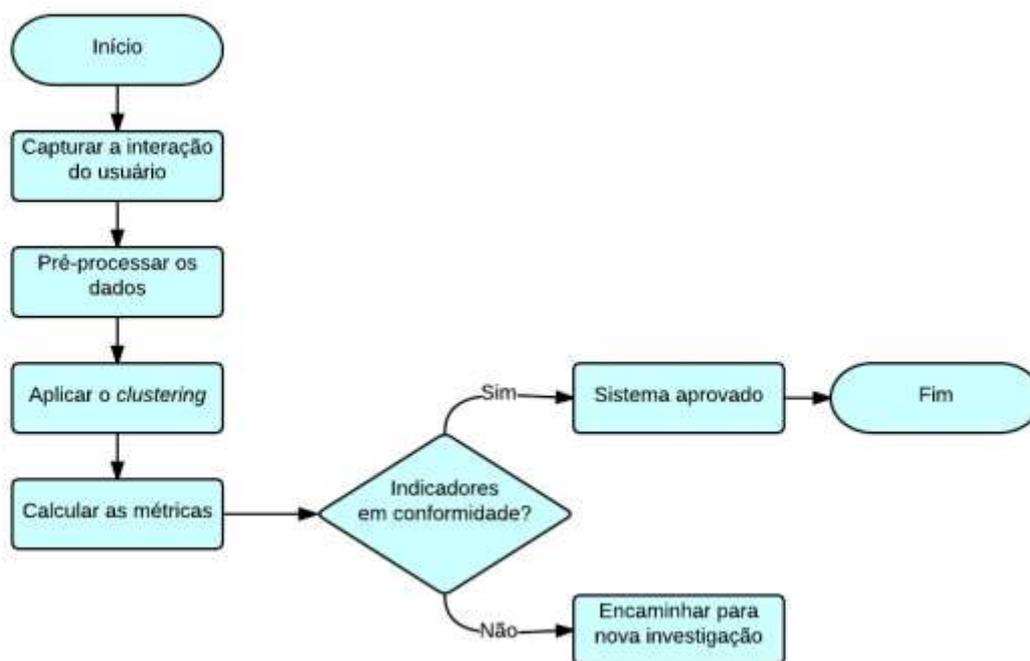


Figura 8 – Estrutura do Projeto de Avaliação de Usabilidade Proposto
Fonte: Autoria Própria

A Figura 8 apresenta a estrutura do processo de avaliação automática de usabilidade proposto. O processo é iniciado com a captura dos dados de interação do usuário. Os dados capturados passam por uma fase de pré-processamento para ficar no formato necessário para a aplicação do algoritmo de *clustering*. As métricas são aplicadas a cada um dos *clusters* gerados para a obtenção dos indicadores para a decisão se o sistema está em conformidade ou existe a necessidade de reprojeção do sistema. Caso a indicação seja reprojeção do sistema, o processo é reiniciado após a adequação para nova avaliação.

4.2.4 Executar a avaliação

Nesta etapa estão previstas as seguintes atividades: obter os valores das métricas; aplicar critérios de decisão para as métricas; aplicar critérios de decisão para a avaliação.

Antes de obter os valores das métricas, o arquivo de *log* deve ser configurado para ficar no formato necessário à aplicação do algoritmo de *clustering*. O algoritmo utiliza cálculos de distância para agrupar os registros semelhantes e, por essa razão, os dados de interação do usuário com o sistema são transformados em dados numéricos conforme

padronização apresentada na Tabela 29. Como o sistema poderia ser acessado utilizando qualquer sistema operacional e qualquer navegador, são definidos números para identificar qual deles o usuário está utilizando. A ocorrência de um evento é representada por 1 (um) e a não ocorrência por 0 (zero). O tempo de conclusão das tarefas é expresso em milissegundos. No caso de usuários que apenas acessaram o sistema sem realizar a tarefa proposta, o tempo é definido em 90000000000. O momento de ocorrência de cada um dos eventos está no formato HHMMSS.

Tabela 29 – Padronização dos dados para execução do agrupamento

Categoria	Dados categóricos arquivo XML	Dados numéricos – arquivo para o agrupamento
Sistema Operacional	Windows	1
	Linux	2
	MacOS	3
	Outros	4
Navegador	Chrome	1
	Firefox	2
	Internet Explorer	3
	Safari	4
	Outros	5
Evento	Ocorrência	1
	Não ocorrência	0
Tempo	Tempo de conclusão da tarefa	Expresso em milissegundos
	Tempo do usuário desistente	90000000000
	Momento de execução do evento	Expresso no formato HHMMSS

Fonte: Autoria Própria.

As entradas do arquivo de *log* original gerado pela WAUTT são transferidas para uma tabela contendo tantas colunas quanto o número de tipos de eventos realizados para cada tarefa e atributos relacionados aos eventos. A quantidade total de colunas existente em cada tabela é determinada pela quantidade de tipos de eventos do usuário com mais interações com o sistema web.

Para os eventos clique, movimento do *mouse* e rolagem também são consideradas a posição do *mouse*, a posição da barra de rolagem e a hora da ocorrência do evento. Por exemplo, no caso de um usuário que executou o clique vinte vezes, serão vinte colunas para o evento clique, mais a mesma quantidade de colunas para a posição *x*, para a posição *y* e para a hora da ocorrência do evento.

A Figura 9 representa um exemplo (fragmento) do arquivo de *log*, em formato XML, gerado pela WAUTT. A sessão do usuário foi identificada pelo IP da máquina (10.20.32.196). O exemplo mostra o horário de início da tarefa (18:30:26) e a ocorrência dos eventos *load* (carregamento da página) e *mousemove* (movimento do *mouse*).

```
<?xml version="1.0"?>
<log applang="PHP" appurl="http://10.20.10.165/dreds/" appname="Teste" appid="2a7725e1a20f84b538e6b8a1e238b95c" tim
type="partial">
- <session timezone="-03" outputtime="2014-03-25T19:07:37.407" inputtime="2014-03-25T18:30:26.023" ip="10.20.32.196" r
- <browser>
  <name>Chrome</name>
  <version>24</version>
</browser>
- <os>
  <name>Windows</name>
</os>
- <navigation timezone="-03" inputtime="2014-03-25T18:30:26.038" id="559" lastupdate="2014-03-25T18:32:35.903">
  <url>http://10.20.10.165/dreds/</url>
  <title>Home</title>
  <width>1366</width>
  <height>667</height>
  - <event timezone="-03" id="46292" time="2014-03-25T18:30:26.035">
    <eventname>load</eventname>
    <tagName>BODY</tagName>
    <text>dreds HomeNew Dred What is DREDS?Just a funny story with actual or fictional characters that you desi
  </event>
  - <event timezone="-03" id="46294" time="2014-03-25T18:30:28.040">
    <eventname>mousemove</eventname>
    <posx>220</posx>
    <posy>131</posy>
    <tagName>P</tagName>
    <text>Just a funny story with actual or fictional characters that you design.</text>
  </event>
```

Figura 9 – Exemplo do arquivo de *log* de eventos gerado pela ferramenta WAUTT
Fonte: Autoria própria.

A Figura 10 representa um exemplo da tabela após o pré-processamento dos dados originais para dados numéricos. Essa tabela está pronta para servir de base para execução do algoritmo X-Means. Nessa tabela constam apenas as informações utilizadas para o cálculo das métricas das subcaracterísticas (da característica de usabilidade).

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Tempo Consumido	Click	Load	MouseDown	MouseMove	MouseUp	PhraseTyped	Scroll	Unload	Navegador	S.O.	Quantidade Total	PosClickX0	PosClickY0	PosClickX1	PosClickY1	PosClickX2	PosClickY2
2	1382000	108	3	113	481	111	3	72	2	1	1	893	0	0	251	183	606	241
3	1779000	85	2	135	425	132	3	36	1	1	1	813	0	0	0	0	523	189
4	1934000	71	2	78	497	78	3	48	1	1	1	718	0	0	0	0	309	166
5	1935000	73	3	99	484	96	3	41	2	1	1	801	590	142	0	0	393	165
6	1937000	104	2	110	406	108	3	45	1	1	1	779	0	0	547	189	261	146
7	1940000	176	2	180	485	180	3	100	1	2	1	1137	0	0	435	164	551	242
8	1942000	158	6	159	409	161	6	33	5	2	1	988	0	0	580	175	400	205
9	1962000	104	2	109	473	109	5	52	1	1	1	855	357	92	383	90	488	138
10	1964000	109	2	120	451	120	3	52	1	1	1	858	689	113	664	135	0	0
11	1967000	161	2	172	598	170	5	58	1	2	1	1163	0	0	341	137	664	140
12	1977000	122	2	108	753	150	2	17	1	1	1	1153	0	0	230	174	290	186
13	2122000	103	2	101	551	100	3	47	1	2	1	908	0	0	514	188	537	208
14	2125000	108	2	106	414	106	2	33	1	1	1	778	0	0	361	188	538	228
15	2134000	134	2	134	583	136	3	88	1	2	1	1071	663	225	663	225	661	225
16	2181000	86	2	90	420	89	2	38	1	1	1	798	853	141	0	0	427	190
17	2184000	110	2	129	464	128	3	43	1	1	1	888	0	0	247	187	631	141
18	2185000	118	2	117	407	117	3	51	1	1	1	818	0	0	268	135	649	135
19	2202000	116	2	119	552	117	3	28	1	1	1	938	448	128	0	0	444	160
20	2203000	129	2	157	728	148	2	140	2	1	1	1309	637	127	0	0	317	142
21	2231000	154	4	170	591	176	3	57	1	1	1	1100	0	0	390	150	230	169
22	2234000	104	2	112	597	111	4	71	1	1	1	942	642	136	0	0	467	161
23	2231000	87	2	89	272	87	3	40	1	1	1	581	0	0	434	187	438	197

Figura 10 – Exemplo do arquivo de *log* de eventos com dados numéricos
Fonte: Autoria própria.

A Figura 10 representa um exemplo da tabela após o pré-processamento dos dados originais para dados numéricos. Essa tabela está pronta para servir de base para execução do algoritmo X-Means. Nessa tabela constam apenas as informações utilizadas para o cálculo das métricas das subcaracterísticas (da característica de usabilidade).

Cada linha da tabela gerada pelo processamento dos dados originais do arquivo XML em dados numéricos representa a ação de um usuário em uma tarefa específica. Cada coluna indica a quantidade de ocorrências ou não de um evento. Para cada usuário também foi calculado o tempo total gasto para finalização da tarefa proposta e foram capturados os detalhes específicos relacionados ao tipo de evento.

As tabelas referentes aos arquivos de *log* gerados pela WAUTT com os dados da captura da interação do usuário na execução das tarefas propostas servem de base para execução do algoritmo de agrupamento X-Means. A execução do agrupamento é realizada com a ferramenta WEKA, versão 3.5.8.

A ferramenta WEKA é um *software* de código aberto que reuni uma coleção de algoritmos de aprendizado de máquina para execução de atividades de mineração de dados (WEKA, 2015). Embora a ferramenta WEKA possua o X-Means implementado, a resposta padrão é quantas instâncias existem em cada grupo (*cluster*). Para obter quais instâncias foram colocadas em cada um dos grupos, foi necessário adaptar o código da ferramenta.

4.2.5 Concluir a avaliação

Esta etapa do processo de avaliação consiste em revisar os dados e apresentar os resultados.

5 AVALIAÇÃO AUTOMÁTICA DE USABILIDADE USANDO *CLUSTERING* E MÉTRICAS

Este capítulo descreve a aplicação da avaliação automática de usabilidade proposta utilizando *clustering* e métricas. Duas avaliações foram realizadas usando diferentes usuários, contextos e sistemas. A Seção 5.1 apresenta a aplicação da avaliação automática com o Sistema para Criação de Diálogos entre Personagens – DREDS e a Seção 5.2 apresenta a aplicação com o Sistema Eletrônico de Administração de Conferências – SEAC.

5.1 AVALIAÇÃO DO SISTEMA DREDS

A primeira avaliação automática de usabilidade usando *clustering* foi realizada com o Sistema para Criação de Diálogos entre Personagens, denominado de DREDS. O objetivo do DREDS consiste em permitir que o usuário elabore uma história e possa compartilhá-la em uma rede social. No DREDS, uma história é composta por um conjunto de diálogos e cada diálogo (ou fala) está associado a um personagem.

O sistema DREDS é uma RIA desenvolvida utilizando o *framework* jQuery (<http://jquery.org>), possui um editor HTML e faz uso de recursos de *drag and drop*, *edit in place* e de caixas de diálogo modais. A interface do DREDS possui uma página inicial, com um pequeno texto descrevendo a aplicação, e a página responsável pela criação da história, mostrada na Figura 11, que é o foco principal da avaliação.

O propósito da avaliação proposta para o DREDS foi de obter indicadores para subcaracterísticas de usabilidade para tomar decisão sobre a aceitação ou não do sistema e sobre a necessidade ou não de investigação com relação a determinados aspectos de usabilidade. As subcaracterísticas de usabilidade avaliadas como requisitos de qualidade são: apreensibilidade, operabilidade e eficiência, proteção ao erro do usuário e estética da interface. Para o DREDS, todas as funcionalidades do sistema foram avaliadas: entrada de dados, alteração da ordem dos diálogos, alteração de personagem e do diálogo, exclusão e formatação.

As métricas para a avaliação são as relacionadas às subcaracterísticas de usabilidade da maneira como foram apresentadas na Seção 4.2.2. Como critério de decisão das métricas nesta avaliação, o valor de 0,75 foi considerado o nível mínimo para a métrica indicar conformidade com a subcaracterística de usabilidade. Esse valor (0,75) foi escolhido de

maneira arbitrária para servir de parâmetro inicial e pode ser ajustado a partir da experiência com a avaliação de mais sistemas (por exemplo, a partir de dados históricos de mais avaliações). Como critério de decisão da avaliação, o sistema pode ser liberado se todas as métricas indicam conformidade, caso contrário o sistema (ou partes do sistema) deve ser encaminhado para nova investigação.

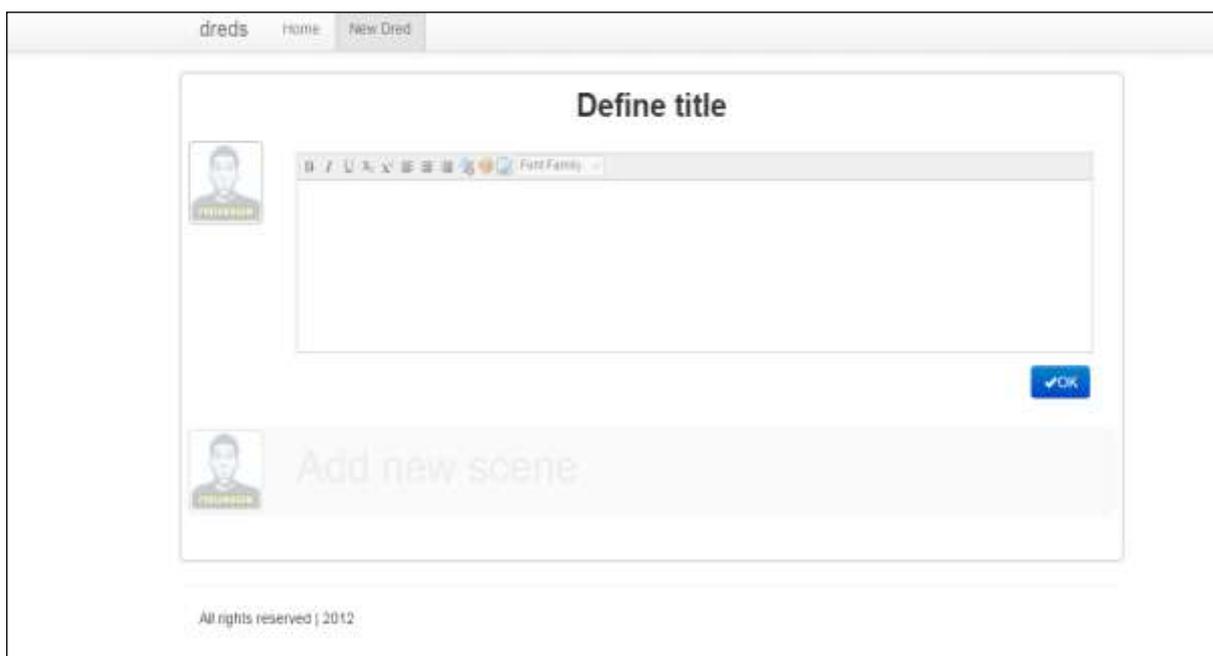


Figura 11 – Página do Sistema DREDS

Fonte: Sistema DREDS

A avaliação do DREDS foi realizada com 36 alunos de um curso da área de Ciência da Computação, da Universidade Tecnológica Federal do Paraná, Câmpus Cornélio Procópio, em um laboratório da graduação e de maneira controlada, isto é, cada tarefa realizada a partir de um mesmo momento. A duração total da avaliação foi de aproximadamente 1 hora e 40 minutos e a participação dos alunos foi voluntária. Cada usuário recebeu um documento com a relação de tarefas que deveria executar. As cinco tarefas propostas foram realizadas em sequência e as tarefas 2, 3, 4 e 5 dependiam da conclusão da primeira tarefa. Embora houvesse a presença de um supervisor durante a aplicação do teste, nenhum usuário recebeu ajuda quanto às tarefas propostas; apenas foram informados dos objetivos, dos dados que seriam capturados e como esses dados seriam analisados.

Para a avaliação do DREDS, foram propostas cinco tarefas (Apêndice B), cada uma com um objetivo específico:

1. Criar a história “As coisas que mais gostamos”, contendo 17 falas, seis personagens e um narrador. Essa tarefa requer que o usuário digite as falas e identifique os personagens corretamente;
2. Introduzir um novo personagem na história. Para essa tarefa, o usuário deve criar três novas falas e inseri-las em posições específicas (necessário utilizar o recurso de *drag and drop*);
3. Alterar a história, trocando dois personagens e suas falas. Nessa tarefa, o usuário poderia ou editar as falas criadas ou excluir as cenas e adicionar novas em substituição;
4. Excluir as falas de dois personagens, que consiste na exclusão dos diálogos em questão;
5. Formatar o texto de algumas falas utilizando os recursos do editor HTML.

No total, com a avaliação do DREDS foram capturados 108.507 eventos, distribuídos entre as cinco tarefas propostas, conforme mostrado no Gráfico 1.

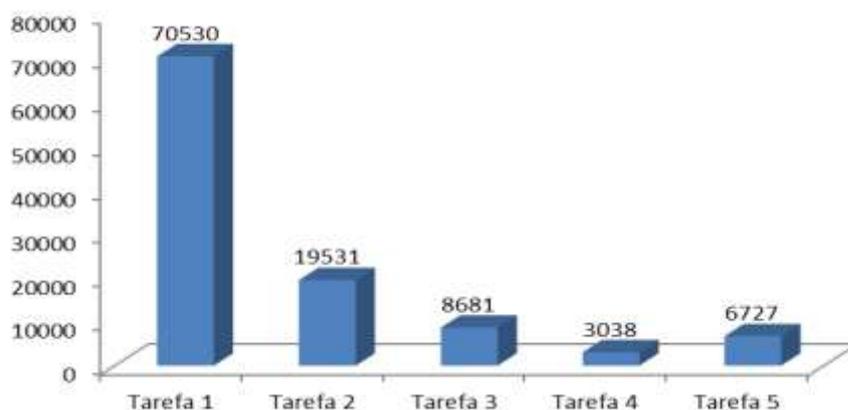


Gráfico 1 – Quantidade de Eventos Capturados por Tarefa
Fonte: Autoria Própria.

Em relação aos tipos de eventos capturados, a utilização do teclado foi o tipo de interação com maior ocorrência, representando 51% (55.339) dos eventos. A distribuição dos eventos por tipo está representada no Gráfico 2.

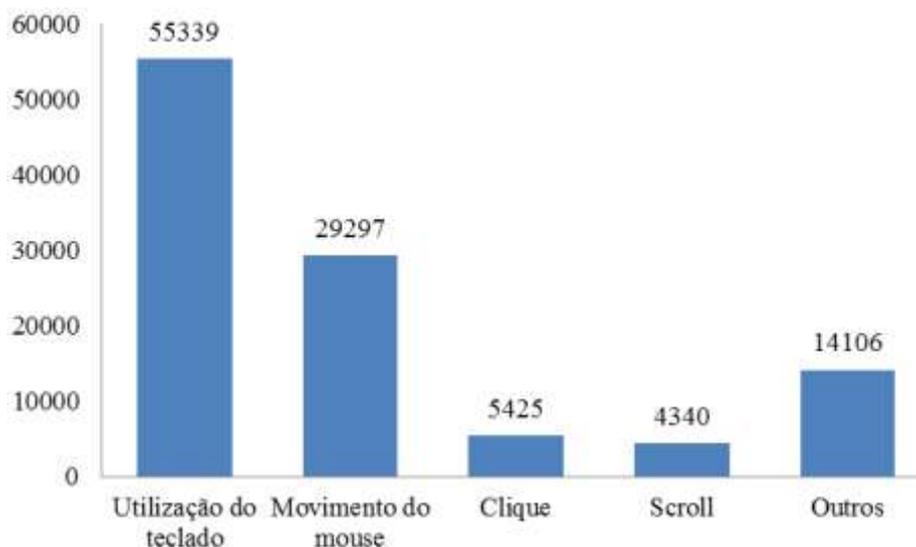


Gráfico 2 – Quantidade de Eventos Capturados por Tipo - DREDS
Fonte: Autoria Própria.

A execução do X-Means nos dados do arquivo de *log* de eventos na primeira Tarefa resultou no agrupamento das instâncias em quatro *clusters*, conforme apresentado na Tabela 30.

Tabela 30 – Execução do X-Means, Tarefa 1, Sistema DREDS

<i>Cluster</i>	Quantidade de instâncias	Instâncias	Características
0	21 (58,33%)	0, 3, 4, 7, 9, 10, 11, 13, 14, 17, 18, 19, 22, 24, 25, 26, 27, 28, 31, 32 e 33	Usuários que mais buscaram por informação para compreender como iriam executar a tarefa.
1	8 (22,22%)	2, 5, 6, 8, 12, 20, 21 e 29	Usuários com maiores valores de apreensibilidade.
2	5 (13,89%)	1, 15, 23, 34 e 35	Usuários que não iniciaram a tarefa (tempo = 90000000) ou desistiram da tarefa.
3	2 (5,56%)	16 e 30	Usuários que fizeram menos operações no sistema.

Fonte: Autoria Própria.

Nas Tabelas 4, 5, e 6 são apresentados os resultados das métricas relacionadas às subcaracterísticas de usabilidade que foram aplicadas aos *clusters* 0, 1 e 3. As instâncias agrupadas no *cluster* 2 representam usuários que não iniciaram ou desistiram da Tarefa 1 e, por esta razão, não foram consideradas para o cálculo das métricas.

Tabela 31 – Métricas do cluster 0, Tarefa 1, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
0	2022857	936	M ₁ – Apreensibilidade	$I_1 = 16/21 \rightarrow I_1 = 0,76$
			M ₂ – Operabilidade/ eficiência	$I_2 = 11/21 \rightarrow I_2 = 0,52$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(15/21) \rightarrow I_3 = 0,29$
			M ₄ – Estética da interface	$I_4 = 1-(((15+14)/2)/21) \rightarrow I_4 = 0,31$

Fonte: Autoria Própria.

Tabela 32- Métricas do cluster 1, Tarefa 1, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
1	1947625	649	M ₁ – Apreensibilidade	$I_1 = 7/8 \rightarrow I_1 = 0,88$
			M ₂ – Operabilidade/ eficiência	$I_2 = 4/8 \rightarrow I_2 = 0,50$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(0/8) \rightarrow I_3 = 1,00$
			M ₄ – Estética da interface	$I_4 = 1-(((0+0)/2)/8) \rightarrow I_4 = 1,00$

Fonte: Autoria Própria.

Tabela 33 – Métricas do cluster 3, tarefa 1, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
3	2013500	441	M ₁ – Apreensibilidade	$I_1 = 1/2 \rightarrow I_1 = 0,50$
			M ₂ – Operabilidade/ eficiência	$I_2 = 1/2 \rightarrow I_2 = 0,50$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(0/2) \rightarrow I_3 = 1,00$
			M ₄ – Estética da interface	$I_4 = 1-(((0+0)/2)/2) \rightarrow I_4 = 1,00$

Fonte: Autoria Própria.

Na Tarefa 1, entrada de dados, a maior parte dos usuários foram agrupados no *cluster* 0 (58,33%). Esses usuários utilizaram mais tempo para concluir a tarefa e os indicadores relacionados à proteção ao erro do usuário ($I_3 = 0,29$) e à estética da interface ($I_4 = 0,31$), foram menos satisfatórios que os resultados encontrados no *cluster* 1, considerando o critério de decisão previamente definido.

Usuários que não buscaram ajuda e, aparentemente, compreenderam melhor a interface do sistema, agrupados no *cluster* 1, não tiveram dificuldade durante a execução da tarefa, o que pode ser constatado pelos valores de proteção ao erro do usuário ($I_3 = 1,00$) e

estética da interface ($I_4 = 1,00$). No entanto, esse *cluster* não reflete a realidade da maioria dos usuários, devido a quantidade de usuários agrupados.

Na segunda tarefa, a execução do X-Means nos dados do arquivo de *log* de eventos resultou no agrupamento das instâncias em três *clusters*, conforme apresentado na Tabela 30.

Tabela 34 – Execução do X-Means, tarefa 2, Sistema DREDS

<i>Cluster</i>	Quantidade de instâncias	Instâncias	Observações
0	9 (25%)	3, 4, 7, 9, 13, 18, 21, 25 e 32	Alto valor de apreensibilidade e valores muito baixos para demais métricas.
1	19 (52,78%)	0, 2, 5, 8, 11, 12, 14, 16, 17, 19, 22, 24, 27, 28, 29, 30, 31 e 33	Baixo valor de operabilidade.
2	8 (22,22%)	1, 6, 10, 15, 23, 26, 34 e 35.	Usuários que não iniciaram a tarefa (tempo = 90000000) ou desistiram.

Fonte: Autoria Própria.

Nas Tabelas 8 e 9 são apresentados os resultados das métricas relacionadas às subcaracterísticas de usabilidade que foram aplicadas aos *clusters* 0 e 1. As instâncias agrupadas no *cluster* 2 representam usuários que não iniciaram ou desistiram da Tarefa 2 e, por esta razão, não foram consideradas para o cálculo das métricas.

Tabela 35 – Métricas do *cluster* 0, Tarefa 2, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
0	2089333	548	M ₁ – Apreensibilidade	$I_1 = 9/9 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 3/9 \rightarrow I_2 = 0,33$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(8/9) \rightarrow I_3 = 0,11$
			M ₄ – Estética da interface	$I_4 = 1-(((9+8)/2)/9) \rightarrow I_4 = 0,05$

Fonte: Autoria Própria.

Tabela 36 – Métricas do *cluster* 1, Tarefa 2, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
1	2038623	311	M ₁ – Apreensibilidade	$I_1 = 18/19 \rightarrow I_1 = 0,95$
			M ₂ – Operabilidade/ eficiência	$I_2 = 10/19 \rightarrow I_2 = 0,53$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(1/19) \rightarrow I_3 = 0,95$
			M ₄ – Estética da interface	$I_4 = 1-(((1+6)/2)/19) \rightarrow I_4 = 0,82$

Fonte: Autoria Própria.

O padrão de comportamento mais comum na Tarefa 2, alteração da ordem dos diálogos, é agrupado no *cluster* 1, representando 52,78% dos usuários, e o valor que sugere maior atenção diz respeito à operabilidade/ eficiência ($I_2 = 0,53$), que está ligado à flexibilidade, ao projeto estético e minimalista e ao controle e liberdade do usuário.

As instâncias agrupadas no *cluster* 0, mesmo tendo uma representatividade menor na tarefa, tiveram valores bem baixos nas métricas operabilidade/ eficiência ($I_2 = 0,33$), proteção ao erro do usuário ($I_3 = 0,11$) e estética da interface ($I_4 = 0,05$).

Na questão da compreensão geral, quantificado pelo valor de apreensibilidade, não houve resultado insatisfatório (menor que 0,75) em nenhum dos *clusters* formados.

Na terceira Tarefa, a execução do X-Means resultou no agrupamento das instâncias em quatro *clusters*, conforme apresentado na Tabela 10.

Tabela 37 – Execução do X-Means, Tarefa 3, Sistema DREDS

<i>Cluster</i>	Quantidade de instâncias	Instâncias	Características
0	9 (25%)	13, 18, 20, 21, 22, 25 e 31	Valores baixos nas métricas de proteção ao erro do usuário e estética da interface
1	9 (25%)	3, 5, 7, 11, 14, 16, 17, 27 e 32	Usuários com maiores valores de apreensibilidade.
2	10 (27,78%)	0, 2, 8, 12, 19, 24, 28, 29, 30 e 33	Usuários com maiores valores em todas as métricas.
3	8 (22,22%)	1, 6, 10, 15, 23, 26, 34 e 35	Usuários que não iniciaram a tarefa (tempo = 90000000) ou desistiram da tarefa.

Fonte: Autoria Própria.

Nas Tabelas 11, 12 e 13 são apresentados os resultados das métricas relacionadas às subcaracterísticas de usabilidade que foram aplicadas aos *clusters* 0, 1 e 2. As instâncias agrupadas no *cluster* 3 representam usuários que não iniciaram ou desistiram da Tarefa 3 e, por esta razão, não foram consideradas para o cálculo das métricas.

Tabela 38 – Métricas do *cluster* 0, Tarefa 3, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
0	2002000	304	M ₁ – Apreensibilidade	$I_1 = 9/9 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 5/9 \rightarrow I_2 = 0,56$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(9/9) \rightarrow I_3 = 0,00$
			M ₄ – Estética da interface	$I_4 = 1-(((9+6)/2)/9) \rightarrow I_4 = 0,17$

Fonte: Autoria Própria.

Tabela 39 – Métricas do *cluster* 1, Tarefa 3, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
1	1947625	649	M ₁ – Apreensibilidade	$I_1 = 9/9 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 2/9 \rightarrow I_2 = 0,22$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(3/9) \rightarrow I_3 = 0,67$
			M ₄ – Estética da interface	$I_4 = 1-(((2+4)/2)/9) \rightarrow I_4 = 0,67$

Fonte: Autoria Própria.

Tabela 40 – Métricas do *cluster* 2, Tarefa 3, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
2	2013500	441	M ₁ – Apreensibilidade	$I_1 = 10/10 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade / eficiência	$I_2 = 1/10 \rightarrow I_2 = 0,60$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(0/10) \rightarrow I_3 = 1,00$
			M ₄ – Estética da interface	$I_4 = 1-(((0+2)/2)/10) \rightarrow I_4 = 0,90$

Fonte: Autoria Própria.

Na Tarefa 3, a quantidade de instâncias em cada *cluster* é semelhante. Analisando o resultado das métricas de operabilidade/ eficiência, proteção ao erro do usuário e estética da interface, infere-se que o algoritmo dividiu as instâncias entre aqueles usuários que tiveram pouca dificuldade em concluir a tarefa, alocados no *cluster* 2, aqueles que tiveram que repetir as operações e passaram parte do tempo apenas navegando, sem executar alguma modificação na tarefa, agrupados no *cluster* 0, e os usuários que tiveram mais dificuldade em relação a operabilidade/ eficiência da interface, *cluster* 1.

Usuários que utilizaram mais tempo navegando pelo sistema antes de executar alguma operação, *cluster* 0, têm valores maiores de operabilidade/ eficiência, o que pode sugerir que a informação de como o usuário deve proceder existe, embora não seja facilmente compreendida pelos os usuários.

Na quarta Tarefa, a execução do X-Means nos dados do arquivo de *log* de eventos resultou no agrupamento das instâncias em quatro *clusters*, conforme apresentado na Tabela 14. As instâncias agrupadas no *cluster* 0 e *cluster* 1 representam usuários que não iniciaram ou desistiram da Tarefa 4 e, por esta razão, não foram consideradas para o cálculo das métricas.

Tabela 41 – Execução do X-Means, Tarefa 4, Sistema DREDS

<i>Cluster</i>	Quantidade de instâncias	Instâncias	Características
0	4 (11,11%)	22, 24, 31 e 33	Usuários que desistiram de concluir a tarefa.
1	8 (22,22%)	1, 6, 10, 15, 23, 26, 34 e 35	Usuários que não iniciaram a tarefa.
2	17 (47,23%)	0, 2, 3, 4, 5, 11, 12, 14, 16, 19, 21, 25, 27, 28, 29, 30 e 32	Usuários com os melhores valores em todas as métricas.
3	7 (19,44%)	7, 8, 9, 13, 17, 18 e 20	Usuários com os menores valores e maior ocorrência de eventos e tempo utilizado.

Fonte: Autoria Própria.

Nas Tabelas 15, 16 e 17 são apresentados os resultados das métricas relacionadas às subcaracterísticas de usabilidade que foram aplicadas aos *clusters* 0, 2 e 3. As instâncias agrupadas no *cluster* 1 representam usuários que não iniciaram ou desistiram da Tarefa 4 e, por esta razão, não foram consideradas para o cálculo das métricas.

Tabela 42 – Métricas do *cluster* 0, Tarefa 4, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
0	1899500	24	M ₁ – Apreensibilidade	$I_1 = 4/4 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 0/4 \rightarrow I_2 = 0,00$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(0/4) \rightarrow I_3 = 0,00$
			M ₄ – Estética da interface	$I_4 = 1-(((0+0)/2)/4) \rightarrow I_4 = 1,00$

Fonte: Autoria Própria.

As instâncias agrupadas no *cluster* 0 referem-se a usuários que desistiram da tarefa, por esse motivo as métricas devem ser desconsideradas da análise. Infere-se que os usuários desistiram da tarefa devido ao baixo número de eventos capturados e pelo tempo reduzido.

Tabela 43 – Métricas do *cluster* 2, Tarefa 4, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
2	2059118	57	M ₁ – Apreensibilidade	$I_1 = 17/17 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 8/17 \rightarrow I_2 = 0,47$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(1/17) \rightarrow I_3 = 0,95$
			M ₄ – Estética da interface	$I_4 = 1-(((1+4)/2)/17) \rightarrow I_4 = 0,85$

Fonte: Autoria Própria.

Tabela 44 – Métricas do cluster 3, Tarefa 4, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
3	2133571	177	M ₁ – Apreensibilidade	$I_1 = 7/7 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 1/7 \rightarrow I_2 = 0,14$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(7/7) \rightarrow I_3 = 0,00$
			M ₄ – Estética da interface	$I_4 = 1-(((7+6)/2)/7) \rightarrow I_4 = 0,07$

Fonte: Autoria Própria.

Na Tarefa 4, funcionalidade de exclusão, o *cluster* 2 é mais representativo (com maior número de instâncias), com 47,23% das instâncias. Nesse *cluster*, a característica que necessita de mais atenção por parte do avaliador é a operabilidade/ eficiência (I₂), que teve um valor de 0,47.

Embora o *cluster* 3 tenha valores baixos para as métricas de operabilidade/ eficiência (I₂), proteção ao erro do usuário (I₃) e estética da interface (I₄), a quantidade de usuários agrupados com esse padrão de comportamento (7 usuários) é menor que o número mínimo de usuários indicado pela Norma ISO/IEC 25030 para ser usado como métrica (8 usuários), o que não impede, no entanto, que essas informações sejam utilizadas, a critério do avaliador.

Na Tarefa 5, a execução do X-Means nos dados do arquivo de *log* de eventos resultou no agrupamento das instâncias em quatro *clusters*, conforme apresentado na Tabela 3018. As instâncias agrupadas no *cluster* 3 representam usuários que não iniciaram ou desistiram da Tarefa 5 e, por esta razão, não foram consideradas para o cálculo das métricas.

Tabela 45 – Execução do X-Means, Tarefa 5, Sistema DREDS

<i>Cluster</i>	Quantidade de instâncias	Instâncias	Características
0	11 (30,56%)	0, 8, 9, 11, 13, 17, 18, 22, 29, 30 e 32	Menores valores de operabilidade/ eficiência, proteção ao erro do usuário e estética da interface.
1	14 (38,89%)	3, 4, 5, 7, 12, 14, 20, 21, 24, 25, 27, 28, 31 e 33	Valores mais satisfatórios em todas as métricas.
2	3 (8,33%)	2, 16 e 19	Usuários que desistiram
3	8 (22,22%)	1, 6, 10, 15, 23, 26, 34 e 35	Usuários que não iniciaram a tarefa.

Fonte: Autoria Própria.

Nas Tabelas 19 e 20 são apresentados os resultados das métricas relacionadas às subcaracterísticas de usabilidade que foram aplicadas aos *clusters* 0 e 1. As instâncias

agrupadas no *cluster 2* e no *cluster 3* representam usuários que não iniciaram ou desistiram da Tarefa 5 e, por esta razão, não foram consideradas para o cálculo das métricas.

Tabela 46 – Métricas do *cluster 0*, Tarefa 5, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
0	2086091	352	M ₁ – Apreensibilidade	$I_1 = 11/11 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 4/11 \rightarrow I_2 = 0,36$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(11/11) \rightarrow I_3 = 0,00$
			M ₄ – Estética da interface	$I_4 = 1-(((11+9)/2)/11) \rightarrow I_4 = 0,09$

Fonte: Autoria Própria.

Tabela 47 – Métricas do *cluster 1*, Tarefa 5, Sistema DREDS

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
1	2021929	242	M ₁ – Apreensibilidade	$I_1 = 14/14 \rightarrow I_1 = 1,00$
			M ₂ – Operabilidade/ eficiência	$I_2 = 8/14 \rightarrow I_2 = 0,57$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(2/14) \rightarrow I_3 = 0,86$
			M ₄ – Estética da interface	$I_4 = 1-(((1+3)/2)/14) \rightarrow I_4 = 0,86$

Fonte: Autoria Própria.

Os padrões de comportamento dos usuários que concluíram a Tarefa 5, alteração de formatação, foram agrupados em dois *clusters*. No *cluster 0* e no *cluster 1*, os usuários tiveram comportamentos opostos, comprovados pelos valores de proteção ao erro do usuário (I_3) e estética da interface (I_4).

Embora a maior quantidade de usuários, agrupados no *cluster 1*, tenha obtido valores maiores, ainda assim esses resultados devem ter mais atenção por parte do avaliador, por se tratar de padrão de comportamento da maioria dos usuários.

5.2 SÍNTESE DOS RESULTADOS: DREDS

A Tabela 48 mostra o resultado da análise dos *clusters* considerados mais significativos de cada tarefa. O *cluster* mais significativo para uma tarefa é aquele com o maior número de instâncias (usuários), que reflete o comportamento mais homogêneo dos

usuários. O objetivo é analisar o padrão mais frequente para a tomada de decisão com relação à métrica da subcaracterística de usabilidade.

Tabela 48 – Síntese das métricas calculadas, Sistema DREDS

Tarefas	Apreensibilidade	Operabilidade/ Eficiência	Proteção ao erro do usuário	Estética da Interface
1. Entrada de dados (<i>Cluster 0</i> : 21 instâncias)	0,76	0,52	0,29	0,31
2. Alteração da ordem dos diálogos (<i>Cluster 1</i> : 19 instâncias)	0,95	0,53	0,95	0,82
3. Alteração da personagem e do diálogo (<i>Cluster 2</i> : 10 instâncias)	1,00	0,60	1,00	0,90
4. Exclusão (<i>Cluster 2</i> : 17 instâncias)	1,00	0,47	0,95	0,85
5. Formatação (<i>Cluster 1</i> : 14 instâncias)	1,00	0,57	0,86	0,86

Fonte: Autoria Própria.

Nota:

Os valores em destaque referem-se às subcaracterísticas e tarefas que não atingiram o valor mínimo de 0,75.

Em relação à subcaracterística apreensibilidade, cuja valor da métrica foi aumentando de acordo com a execução das tarefas, condiz com o conceito da subcaracterística, que reflete a maneira que o usuário se comporta da primeira vez que acessa o sistema ou quando não tem muita experiência com a interface.

Os menores valores (em negrito), que sugerem mais atenção do avaliador, dizem respeito à subcaracterística operabilidade/ eficiência. Esse resultado significa que os usuários não mantiveram um comportamento condizente com um sistema que fosse fácil de aprender, utilizar e controlar. Outro fato que pode corroborar essa característica é a quantidade de usuários que não conseguiram concluir todas as tarefas propostas. Dos 36 usuários que participaram da avaliação, 11 não completaram todas as tarefas propostas (30,56%). Essa informação foi obtida a partir da análise manual dos usuários agrupados no *Cluster 2* da Tarefa 1 (2 usuários), *Cluster 2* da Tarefa 2 (2 usuários), *Cluster 3* da Tarefa 3 (8 usuários), *Cluster 1* da Tarefa 4 (8 usuários) e *Cluster 3* da Tarefa 5 (8 usuários).

Considerando as subcaracterísticas proteção ao erro do usuário e estética da interface, as métricas que merecem mais atenção do avaliador são as relacionadas com a entrada de dados (Tarefa 1). Nessa tarefa, a maior parte dos usuários precisou repetir operações no sistema para conseguir concluir a tarefa proposta, além de passar algum tempo apenas navegando pela interface, sem executar nenhuma operação. O tempo utilizado apenas para navegação pela interface sugere que os usuários estão procurando por informações antes de decidir qual o próximo passo.

Considerando o critério de decisão das métricas e o critério de decisão para avaliação, definidos na Seção 4.2.2, a subcaracterística de operabilidade/ eficiência não atingiu o valor mínimo 0,75 nas cinco tarefas propostas para o DREDS. Em relação à proteção ao erro do usuário e estética da interface os indicadores obtidos pelo cálculo das métricas também não está em conformidade para a Tarefa 1, entrada de dados. O resultado da avaliação automática de usabilidade sugere que o sistema seja encaminhado para nova investigação, com foco nas funcionalidades que não estão em conformidade com os critérios previamente definidos.

5.3 AVALIAÇÃO DO SISTEMA SEAC

O Sistema Eletrônico de Administração de Conferências, denominado de SEAC, foi utilizado para realizar a segunda avaliação automática de usabilidade usando *clustering* e métricas. O objetivo do sistema é auxiliar nos processos de planejamento, organização e gestão de conferências. Algumas das funções do sistema são: edição e divulgação de chamadas de trabalhos, gerenciamento do processo de chamadas de trabalhos, cadastro de participantes, organização de fóruns e submissão de arquivos (trabalhos). O SEAC foi desenvolvido com base no *software* livre *Open Conference Systems* (OCS) pela Coordenadoria de Gestão da Tecnologia da Informação (COGETI) da Universidade Tecnológica Federal do Paraná (UTFPR), Câmpus Cornélio Procópio. O SEAC foi escolhido porque está em uso pela COGETI para o gerenciamento de eventos do câmpus.

Assim como na avaliação anterior, o propósito da avaliação proposta para o SEAC é obter indicadores para as subcaracterísticas de usabilidade para tomar decisão sobre a aceitação ou não do sistema e sobre a necessidade ou não de investigação com relação a determinados aspectos de usabilidade. As subcaracterísticas de usabilidade avaliadas como

requisitos de qualidade são: apreensibilidade, operabilidade e eficiência, proteção ao erro do usuário e estética da interface.

Para o critério de decisão das métricas foi definido o valor de 0,75, como indicador mínimo de conformidade com a subcaracterística de usabilidade. Como critério de decisão da avaliação, o sistema pode ser liberado se todas as métricas indicam conformidade, caso contrário o sistema (ou partes do sistema) deve ser encaminhado para reprojeção.

A avaliação do SEAC foi realizada com 41 servidores da Universidade Tecnológica Federal do Paraná, Câmpus Cornélio Procópio. Cada usuário (servidor) realizou a tarefa proposta no seu próprio ambiente de trabalho. A captura dos dados foi feita em dois dias. Os usuários receberam um documento explicando a tarefa que deveriam fazer (Apêndice C). A tarefa consistia no envio de um arquivo de escolha do usuário utilizando o SEAC. Para enviar o arquivo era necessário se cadastrar previamente no sistema. Como não houve supervisão durante a execução do processo e não havia a possibilidade de separar a tarefa de cadastro da tarefa de submissão de arquivos, a ferramenta WAUTT gerou um único arquivo de *log* com dados dos eventos capturados.

A interface inicial do SEAC é apresentada na Figura 12. As demais interfaces, obrigatórias para conclusão da tarefa proposta aos usuários, constam no Anexo D.

Para a tarefa proposta com o Sistema SEAC foram capturados 15.701 eventos com a interação do usuário. Em relação ao tipo de evento capturado, o movimento do *mouse* foi o tipo de evento com maior ocorrência, representando 63% dos eventos. A distribuição dos eventos por tipo está representada no Gráfico 3.



Figura 12 – Página de acesso ao SEAC

Fonte: Sistema Eletrônico de Administração de Conferências.

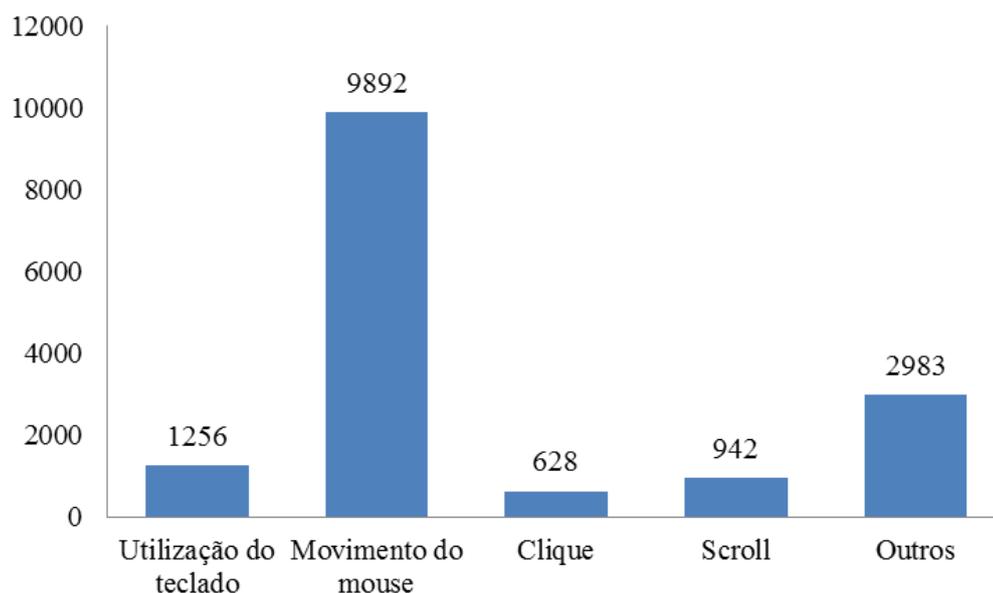


Gráfico 3 – Quantidade de Eventos Capturados por Tipo – Sistema SEAC
Fonte: Autoria Própria.

A execução do X-Means nos dados do arquivo de *log* de eventos da Tarefa proposta resultou no agrupamento das instâncias em três *clusters*, conforme apresentado na Tabela 22.

Tabela 49 – Execução do X-Means , Sistema SEAC

<i>Cluster</i>	Quantidade de instâncias	Instâncias	Características
0	4 (9,76%)	8, 14, 33 e 38	Usuários que utilizaram mais tempo e mais eventos para concluir a tarefa.
1	22 (53,66%)	0, 4, 5, 7, 13, 16, 19, 21, 22, 23, 24, 25, 27, 28, 30, 31, 32, 34, 35, 36, 39 e 40	Todos os usuários desse <i>cluster</i> cometeram mais erros que a média geral.
2	15 (36,59%)	1, 2, 3, 6, 9, 10, 11, 12, 15, 17, 18, 20, 26, 29 e 37	Usuários que gastaram mais tempo para concluir a tarefa, no entanto cometeram menos erros.

Fonte: Autoria Própria.

Nas Tabelas 23 e 24 são apresentados os resultados das métricas relacionadas às subcaracterísticas de usabilidade que foram aplicadas aos *clusters*. As instâncias agrupadas no *cluster* 0 tem pouca representatividade, de acordo com os parâmetros definidos pela ISO/IEC 25030 (mínimo de oito usuários para testes de usabilidade com usuários) e por esse motivo foram desconsiderados da análise.

Tabela 50 – Métricas do *cluster 1*, Sistema SEAC

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
1	2215955	276	M ₁ – Apreensibilidade	$I_1 = 12/22 \rightarrow I_1 = 0,55$
			M ₂ – Operabilidade/ eficiência	$I_2 = 16/22 \rightarrow I_2 = 0,73$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(0/22) \rightarrow I_3 = 0,00$
			M ₄ – Estética da interface	$I_4 = 1-(((0+4)/2)/22) \rightarrow I_4 = 0,91$

Fonte: Autoria Própria.

Tabela 51 – Métricas do *cluster 2*, Sistema SEAC

<i>Cluster</i>	Tempo médio (ms)	Média de eventos	Métrica	Resultado
2	2699467	452	M ₁ – Apreensibilidade	$I_1 = 10/15 \rightarrow I_1 = 0,67$
			M ₂ – Operabilidade/ eficiência	$I_2 = 10/15 \rightarrow I_2 = 0,67$
			M ₃ – Proteção ao erro do usuário	$I_3 = 1-(11/15) \rightarrow I_3 = 0,27$
			M ₄ – Estética da interface	$I_4 = 1-(((13+5)/2)/15) \rightarrow I_4 = 0,40$

Fonte: Autoria Própria.

Na maioria das instâncias agrupadas no *cluster 1*, os usuários tiveram que repetir procedimentos no sistema, motivo pelo qual o valor de proteção ao erro do usuário (I_3) chegou a 0,00, implicando em grande atenção por parte do avaliador. Embora esse índice não tenha zerado no *cluster 2*, o valor ainda continuou ainda muito baixo (0,27).

Para o *cluster 1*, com maior representatividade 53,66% dos usuários, as métricas de operabilidade/ eficiência ($I_2=0,73$) e estética da interface ($I_4=0,91$) tiveram valores mais satisfatórios que os demais usuários, ficando a critério do avaliador o limite que deve ser considerado crítico e quantidade de usuários mínima que exige mais atenção.

5.4 SÍNTESE DOS RESULTADOS: SEAC

A Tabela 52 mostra o resultado da análise do *cluster* considerado mais significativo. A aplicação do X-Means agrupou as instâncias em três *clusters*. Como os usuários que participaram da avaliação realizaram uma única tarefa, somente um cluster é apresentado nessa tabela (aquele com maior número de instâncias).

Tabela 52 – Síntese das métricas calculadas, Sistema SEAC

	Apreensibilidade	Operabilidade/ Eficiência	Proteção ao erro do usuário	Estética da Interface
<i>Cluster 1 (22 instâncias)</i>	0,55	0,73	0,00	0,91

Fonte: Autoria Própria.

Nota:

O valores em destaque referem-se as subcaracterísticas e tarefas que não atingiram o valor mínimo de 0,75.

Os indicadores calculados pelas métricas de apreensibilidade, operabilidade/ eficiência e proteção ao erro do usuário (em destaque) não atingiram o valor mínimo definido como critério de decisão para as métricas (0,75), descrito na Seção 4.2.2. O valor de 0,00 para a métrica de proteção ao erro do usuário indica que todos os usuários agrupados no *cluster* mais significativo precisaram repetir operações no sistema para conseguir concluir a tarefa proposta. O resultado da avaliação automática de usabilidade sugere que o sistema seja encaminhado para nova investigação.

6 COMPARAÇÃO ENTRE AVALIAÇÃO AUTOMÁTICA DE USABILIDADE E AVALIAÇÃO HEURÍSTICA

Para comparar os resultados obtidos com a avaliação automática de usabilidade usando análise de *log* com *clustering* e métricas, uma avaliação heurística foi realizada com os dois sistemas avaliados: DREDS e SEAC. O objetivo dessa comparação foi verificar se os problemas identificados pela avaliação heurística estão relacionados às subcaracterísticas de usabilidade cuja métrica não está em conformidade. A relação entre problema e subcaracterística é estabelecida pela correspondência entre as heurísticas violadas pelos problemas encontrados e as heurísticas relacionadas às subcaracterísticas de usabilidade. A Seção 6.1 descreve a avaliação heurística do Sistema DREDS e a Seção 6.2 descreve a avaliação heurística do Sistema SEAC. As comparações da avaliação automática de usabilidade com a avaliação heurística com relação ao DREDS e ao SEAC são apresentadas nas Seções 6.4 e 6.5, respectivamente. As limitações da avaliação automática de usabilidade proposta são descritas na Seção 6.6.

6.1 AVALIAÇÃO HEURÍSTICA DO SISTEMA DREDS

Para o DREDS, a avaliação heurística foi realizada por 11 estudantes de um curso de graduação da área de Ciência da Computação que estavam cursando a disciplina de Interação Homem-Computador. A avaliação heurística pode ser realizada por um único avaliador, mas o emprego de três a cinco pessoas na avaliação pode identificar 65-75% dos problemas de usabilidade, conforme análises de custo-benefício (NIELSEN, 1994). A avaliação foi planejada para envolver 11 avaliadores para maximizar a quantidade de problemas encontrados e compensar a possível falta de experiência no método de avaliação.

O resultado da avaliação heurística do DREDS, agrupado por heurística violada, de acordo com a opinião dos especialistas é apresentado no Apêndice E. Os resultados foram agrupados por heurística para facilitar a comparação da informação com o resultado da avaliação automática de usabilidade. Nesse caso, um problema pode aparecer repetido quando está associado a mais de uma heurística, ao contrário de um relatório de problemas de usabilidade tradicionalmente utilizado em uma avaliação heurística.

Foram encontrados cinquenta e nove problemas de usabilidade em nove das dez heurísticas consideradas pelos avaliadores. Na Tabela 53 é apresentada a quantidade de problemas de usabilidade encontrados conforme a heurística violada e nível de severidade.

Tabela 53 – Quantidade de problemas encontrados na avaliação heurística do DREDS

Nº	Heurística violada	Nível de Severidade				
		1	2	3	4	Total
		Quantidade de Problemas				
1	Visibilidade do estado atual do sistema		3	1	-	4
2	Correlação entre o sistema e o mundo real	1	1	-	-	2
3	Controle e liberdade do usuário	1	5	5	-	11
4	Consistência e padrões	1	3	3	-	7
5	Prevenção de erros	-	7	6	-	13
6	Reconhecimento ao invés de memorização	1	1	-	-	2
7	Flexibilidade e eficiência de uso	-	7	2	-	9
8	Projeto estético e minimalista	1	3	3	-	7
9	Suporte ao usuário no reconhecimento de erros	-	-	-	-	-
10	Informações de ajuda e documentação	2	2	-	-	4
	Total	7	32	20	-	59

Fonte: Autoria Própria.

A Tabela 53 mostra que 33,90% dos problemas de usabilidade encontrados pelos especialistas na avaliação heurística do DREDS são considerados graves, 54,24% são considerados simples e os outros 11,86% são problemas cosméticos. 79,66% dos problemas estão concentrados em cinco das dez heurísticas: prevenção de erros, controle e liberdade do usuário, flexibilidade e eficiência de uso, projeto estético e minimalista.

6.2 AVALIAÇÃO HEURÍSTICA DO SISTEMA SEAC

Para o SEAC, a avaliação heurística foi realizada por três avaliadores especialistas que são professores da disciplina Interação Homem-Computador de cursos de graduação e mestrado da área de Ciência da Computação. O Apêndice F apresenta o resultado da avaliação heurística do SEAC, agrupado por heurística violada. Os resultados foram agrupados por heurística para facilitar a comparação da informação com o resultado da avaliação automática de usabilidade.

Na avaliação heurística do SEAC, foram encontrados trinta e sete problemas de usabilidade nas dez heurísticas consideradas pelos avaliadores. Na Tabela 54 é apresentada a quantidade de problemas de usabilidade encontrados conforme a heurística violada e nível de severidade.

Tabela 54 – Quantidade de problemas encontrados na avaliação heurística do SEAC

Nº	Heurística violada	Nível de Severidade				
		1	2	3	4	Total
		Quantidade de Problemas				
1	Visibilidade do estado atual do sistema	1	2	3	2	8
2	Correlação entre o sistema e o mundo real	-	1	2	2	5
3	Controle e liberdade do usuário	-	-	-	2	2
4	Consistência e padrões	-	1	2	-	3
5	Prevenção de erros	-	1	1	2	4
6	Reconhecimento ao invés de memorização	1	-	1	-	2
7	Flexibilidade e eficiência de uso	-	-	2	-	2
8	Projeto estético e minimalista	-	1	3	-	4
9	Suporte ao usuário no reconhecimento de erros	-	-	2	-	2
10	Informações de ajuda e documentação	-	2	2	1	5
	Total	2	8	18	9	37

Fonte: Autoria Própria.

A Tabela 54 mostra que 48,65% dos problemas de usabilidade encontrados pelos especialistas na avaliação heurística do SEAC são considerados graves, 24,32% são considerados catastróficos, 21,62% são considerados simples e os outros 5,41% são problemas cosméticos. Do total de problemas, 70,27% estão concentrados em cinco das dez heurísticas: visibilidade do estado atual do sistema, correlação entre o sistema e o mundo real, prevenção de erros, projeto estético e minimalista e informações de ajuda e documentação.

6.3 COMPARAÇÃO: SISTEMA DREDS

O Quadro 4 relaciona os resultados das duas avaliações realizadas com o DREDS, apresentando as métricas calculadas para as subcaracterísticas de usabilidade e a quantidade de problemas identificados de acordo com as heurísticas violadas.

O resultado esperado pela comparação dos métodos de avaliação era que a relação entre o indicador resultante da aplicação das métricas aos *clusters* na avaliação automática e a

quantidade de problemas encontrados pela avaliação heurística fosse inversamente proporcional, ou seja, quanto menor o indicador resultante da avaliação automática, maior o número de problemas encontrados pelos avaliadores.

Para a subcaracterística de apreensibilidade, que teve o maior indicador na avaliação automática (0,94), os avaliadores encontraram 13 problemas, o que equivale a 22% do total dos problemas. Em relação à subcaracterística de operabilidade/ eficiência, a avaliação automática resultou em um indicador de 0,54, o menor para as quatro métricas, e os avaliadores encontraram a maior quantidade de problemas na avaliação heurística, o que equivale a 45,76% do total dos problemas.

Comparando as outras duas subcaracterísticas, proteção ao erro do usuário e estética da interface, os valores dos indicadores obtidos na avaliação automática indicam conformidade (0,81 e 0,75) para estas subcaracterísticas. No entanto, os valores estão muito próximos do valor limite que foi definido como critério (0,75). O número de problemas de usabilidade identificados na avaliação heurística relacionados a essas subcaracterísticas pode estar indicando que o valor limite não é adequado para a tomada de decisão. Nesse sentido, em um ambiente de produção de software, o valor limite pode ser definido considerando os valores históricos de avaliações realizadas em projetos anteriores.

Subcaracterística de Usabilidade	Heurísticas de Usabilidade	Métrica	Nº de problemas	
Apreensibilidade	Correlação entre o sistema e o mundo real	0,94	2	13
	Reconhecimento ao invés de memorização		2	
	Flexibilidade e eficiência de uso		9	
Operabilidade/ Eficiência	Controle e liberdade do usuário	0,54	11	27
	Flexibilidade e eficiência de uso		9	
	Projeto estético e minimalista		7	
Proteção ao erro do usuário	Visibilidade do estado atual do sistema	0,81	4	17
	Prevenção de erro		13	
Estética da Interface	Projeto estético e minimalista	0,75	7	16
	Flexibilidade e eficiência de uso		9	

Quadro 4 – Comparação entre as Avaliações do DREDS

Fonte: Autoria própria.

6.4 COMPARAÇÃO: SISTEMA SEAC

O Quadro 5 relaciona os resultados das duas avaliações realizadas com o SEAC, apresentando as métricas calculadas para as subcaracterísticas de usabilidade e a quantidade de problemas identificados de acordo com as heurísticas violadas.

Subcaracterística de Usabilidade	Heurísticas de Usabilidade	Métrica	Nº de problemas	
Apreensibilidade	Correlação entre o sistema e o mundo real	0,55	5	9
	Reconhecimento ao invés de memorização		2	
	Flexibilidade e eficiência de uso		2	
Operabilidade/ Eficiência	Controle e liberdade do usuário	0,73	2	8
	Flexibilidade e eficiência de uso		2	
	Projeto estético e minimalista		4	
Proteção ao erro do usuário	Visibilidade do estado atual do sistema	0,00	8	12
	Prevenção de erro		4	
Estética da Interface	Projeto estético e minimalista	0,91	4	6
	Flexibilidade e eficiência de uso		2	

Quadro 5 – Comparação entre as Avaliações do SEAC

Fonte: Autoria própria.

Na comparação do método automático com a avaliação heurística para o SEAC, também foi possível verificar que quanto maior o valor do indicador resultante da aplicação das métricas, menor o número de problemas encontrados na avaliação heurística. Essa relação foi observada nas subcaracterísticas de apreensibilidade e proteção ao erro do usuário.

Em relação à subcaracterística de estética da interface também é possível notar relação semelhante. Para esta subcaracterística foi obtido o maior valor de indicador entre as métricas (0,91) e encontrado o menor número de problemas pelos avaliadores (6 problemas, equivalente a 16% do total dos problemas).

Para a subcaracterística de proteção ao erro do usuário, cujo indicador teve o menor valor de todas as métricas (0,00), foi encontrado o maior número de problemas pelos avaliadores. Dos 37 problemas encontrados, 32% estão divididos entre as heurísticas relacionadas à métrica de proteção ao erro do usuário.

O indicador (0,73) relacionado à métrica de operabilidade/ eficiência e o indicador (0,55) relacionado à métrica da apreensibilidade, estão coerentes com a quantidade de

problemas encontrados na avaliação heurística, 8 (oito) e 9 (nove) problemas, respectivamente.

6.5 LIMITAÇÕES

As fragilidades encontradas com a execução da avaliação automática de usabilidade proposta neste trabalho podem ser descritas de acordo com os itens a seguir:

- Definição dos critérios de decisão para métricas e critérios de decisão para avaliação. Em geral, a área de Engenharia de Software tem enfrentado dificuldades com a falta de valores de referência para comparação. A maior parte das atividades relacionadas a métricas de *software* não fornecem informações para a tomada de decisão (LIMA; RESENDE, 2012). Assim, os valores escolhidos podem não ser adequados para qualquer sistema ou contexto de avaliação automática de usabilidade.
- Especificação do problema de usabilidade encontrado. Embora a abordagem indique uma funcionalidade com problemas relacionados a algumas subcaracterísticas de usabilidade, não é possível conhecer especificamente qual problema precisa ser corrigido. Uma alternativa para essa limitação seria aplicar outros métodos de avaliação de usabilidade para aprofundar a investigação, como o protocolo *think aloud*.
- Obtenção de dados quantitativos através dos arquivos de *log* de eventos que possibilitem a criação de métricas. Como a abordagem proposta utiliza dados numéricos, nem todas as subcaracterísticas de usabilidade podem ser mensuradas através da interpretação dos dados contidos no arquivo de *log* de eventos. Por essa razão, dentre as dez subcaracterísticas de usabilidade previstas pela Série SQuaRe, apenas cinco foram consideradas nos cálculos das métricas (apreensibilidade, operabilidade, eficiência, proteção ao erro do usuário e estética da interface).

Algumas sugestões para dar continuidade à avaliação automática de usabilidade proposta e minimizar os problemas encontrados com as limitações acima estão descritos na próxima Seção.

7 CONCLUSÕES

Neste trabalho foi apresentada uma abordagem para a avaliação automática de usabilidade de RIAs com base na aplicação de *clustering* e métricas à análise de *log*. A utilização de informação no nível de elementos da página e de eventos associados com a interação do usuário com RIAs como entrada para processos automáticos tem potencial para obter resultados para reduzir o esforço do avaliador e o tempo de duração do processo e, conseqüentemente, reduzir também o custo.

A abordagem proposta neste trabalho foi utilizada na avaliação de dois sistemas web diferentes: DREDS e SEAC. As avaliações desses sistemas foram realizadas com usuários e cenários distintos. A preparação necessária para realizar cada uma dessas avaliações (e a posterior realização) permitiu a generalização de alguns aspectos do processo de avaliação com base na instanciamento do *framework* para a avaliação da qualidade de produto de *software* proposto na série SQuARE. Nesse sentido, outras técnicas ou algoritmos de *clustering* e novas métricas ou alterações das que foram utilizadas podem ser incorporadas ao processo apresentado no Capítulo 4.

Os indicadores obtidos na avaliação automática foram comparados com os resultados da avaliação heurística dos dois sistemas web. A realização da avaliação dos sistemas web também com um método de avaliação de usabilidade tradicional foi utilizada para a validação da abordagem proposta, uma vez que foram identificados problemas de usabilidade relacionados às mesmas subcaracterísticas de usabilidade consideradas. As heurísticas associadas às subcaracterísticas foram utilizadas para estabelecer uma relação entre indicadores e problemas de usabilidade associados à heurística violada.

O *clustering* foi uma técnica útil para reduzir o volume de dados e centrar a tomada de decisão a partir da análise apenas dos *clusters* considerados significativos. No entanto, essa não é a única estratégia que pode ser adotada, pois avaliador também pode estar interessado nos comportamentos menos frequentes ou que ocorrem de maneira isolada. Em qualquer caso, o uso de *clustering* pode reduzir o tempo de realização da avaliação de usabilidade, uma vez que o avaliador pode concentrar os seus esforços em determinados grupos de usuários com comportamentos semelhantes.

A aplicação de métricas aos *clusters* permitiu obter indicadores para tomada de decisão com relação a algumas das subcaracterísticas de usabilidade. Além de fornecer uma maneira de obter uma avaliação de usabilidade quantitativa, os indicadores podem ser utilizados como base para a comparação antes e após alguma alteração na interface do sistema

web. Dessa maneira, a avaliação automática de usabilidade realizada como a abordagem proposta tem potencial para auxiliar no desenvolvimento de sistemas web.

7.1 CONTRIBUIÇÕES

As principais contribuições deste trabalho podem ser destacadas de acordo com os seguintes itens:

- Definição de uma abordagem para avaliação automática de usabilidade de RIAs por meio de análise de *log* com *clustering* e métricas e utilizando dados no nível de elementos da página e eventos relacionados à interação do usuário.
- Definição de métricas que fornecem indicadores quantitativos em relação à usabilidade de RIAs, ao invés das informações subjetivas apresentadas com a análise realizada por avaliadores com apresentação de relatórios exclusivamente descritivos. A definição dessas métricas considerou as subcaracterísticas de usabilidade: apreensibilidade, operabilidade, eficiência, proteção ao erro do usuário e estética da interface.

7.2 TRABALHOS FUTUROS

Alguns dos trabalhos futuros que podem dar continuidade às contribuições apresentadas são resumidos a seguir:

- A abordagem proposta para avaliação automática de usabilidade foi aplicada em RIAs com vários problemas de usabilidade. Dessa maneira, era esperado que os indicadores apresentassem valores que seriam considerados baixos para algumas das subcaracterísticas de usabilidade. A avaliação automática de usabilidade da maneira proposta deve ser aplicada também em RIAs consideradas de qualidade em termos de usabilidade.
- A comparação dos indicadores obtidos antes e após alguma alteração na interface de RIAs. Assim, seria possível verificar se realmente houve melhoria depois de realizada a alteração. Além disso, também seria possível estimar o esforço em termos de custo para corrigir problemas relacionados às subcaracterísticas de usabilidade.

- A definição de métricas específicas para a acessibilidade, que também é uma subcaracterística de usabilidade. Nesse sentido, existem diversas diretrizes e recomendações para acessibilidade que poderiam servir como base para a definição de métricas.

Aplicação de outros métodos de teste com usuários nos sistemas avaliados, como o protocolo *think aloud*, para realizar a validação dos resultados da avaliação automática de usabilidade proposta. Neste trabalho, a proposta de avaliação automática de usabilidade com base em análise de log, tipicamente um método de teste com usuários, foi comparada com a avaliação heurística, tipicamente um método de inspeção. Portanto, um trabalho necessário seria realizar uma comparação com outros métodos de teste com usuários.

REFERÊNCIAS

- ALVES, R.; VALENTE, P.; NUNES, N. J.. The state of user experience evaluation practice. In: **PROCEEDINGS OF THE 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational**. ACM, 2014. p. 93-102.
- AMALFITANO, D.; FASOLINO, A. R.; POLCARO, A.; TRAMONTANA, P. Comprehending Ajax web applications by the DynARIA tool. In: **PROCEEDINGS OF THE Conference on the Quality of Information and Communications Technology (QUATIC)**, 2010, Porto, p. 122-131.
- ATTERER, R; SCHIMIDT, A. Tracking the Interaction of Users with AJAX Applications for Usability Testing. In: **PROCEEDINGS OF THE SIGCHI Conference on Human Factors in Computing Systems**, 2007, California, p. 1347-1350.
- ASHRAF, T.; ASHRAF, I. A summarized review on web usage mining. **International Journal of Medical Informatics**, 2014, v. 4, n. 4, p. 2736–2739.
- BALBO, S., GOSCHNICK, S., TONG, D., PARIS, C., Leading Usability Evaluations to WAUTER. In: **PROCEEDINGS OF THE Eleventh Australian World Wide Web Conference**, Gold Coast, Australia, 2005.
- BASTIEN, J. C. Usability testing: a review of some methodological and technical aspects of the method. **International Journal of Medical Informatics**, v. 79, n. 4, p. e18–e23.
- BLANDFORD, A.; KEITH, S.; CONNELL, I.; EDWARDS, H. Analytical Usability Evaluation for Digital Libraries: A Case Study. In: **PROCEEDINGS OF THE 2004 Joint ACM/IEEE Conference on Digital Libraries**, 2004. Tucson: ACM Press, 2004. p. 27-36.
- CARTA, T.; PATERNÒ, F.; DE SANTANA, V. Web usability probe: a tool for supporting remote usability evaluation of web sites. In: **PROCEEDINGS OF THE Human-Computer Interaction, 2011**. Springer Berlin Heidelberg, 2011. p. 349-357.
- CONTE, T., MASSOLAR, J., MENDES, E., TRAVASSOS, P. G. H. Web Usability Inspection Technique Based on Design Perspectives. **IET Software Journal**, Special Issue: Selected Papers, 2007, p.1-18.
- CYBIS, W.; BETIOL, A. H.; FAUST, R. **Ergonomia e usabilidade: conhecimentos, métodos e aplicações**. 2 ed. São Paulo: Novatec, 2010.
- DIAS, C. **Usabilidade na web: criando portais mais acessíveis**. 2 ed. Rio de Janeiro: Alta Books, 2007.
- DUAN, J. L.; LIU, S. X. Application on web mining for web usability analysis. In: **PROCEEDINGS OF THE International Conference on Machine Learning and Cybernetics**, 2012, Xian, China, p. 1981–1985.
- EVERITT, B.; LANDAU, S.; LEESE, M.; STAHL, D. **Cluster Analysis**. 5 ed. United Kingdom: John Wiley & Sons, Ltd, 2011.

FERNANDEZ, A.; INSFRAN, E.; ABRAHÃO, S.. Usability evaluation methods for the web: A systematic mapping study. **Information and Software Technology**, v. 53, n. 8, 2011, p. 789-817.

FREITAS, R. C.; DUTRA, M. DE A. Usabilidade e interatividade em sistemas web para cursos online. **Revista Brasileira de Informática na Educação**, v. 17, n. 2, 2009, p. 48.

HAN, J.; KAMBER, M. **Data mining concepts and techniques**. 3 ed. Waltham: Morgan Kaufmann, 2011.

HOLLINGSSED, T.; NOVICK, D. G. Usability inspection methods after 15 years of research and practice. In: **PROCEEDINGS OF THE 25th annual ACM international conference on design of communication**. ACM, 2007. p. 249-255.

INTERNATIONAL STANDART ORGANIZATION/ INTERNATIONAL ELECTROTECHNICAL COMMISSION. **ISO/IEC 9126**: Information technology - Software product evaluation - Quality characteristics and guidelines for their use. Geneva, 1991.

_____. **ISO/IEC 9126-1**: Software Engineering – Product quality - Part 1. Geneva, 2000.

_____. **ISO 25000 Software Product Quality**. Disponível em: <<http://iso25000.com/index.php/en/>>. Acesso em: 23 jan. 2015.

_____. **ISO/IEC 25010**: System and Software engineering - System and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models. Geneva, 2011.

_____. **ISO/IEC 25020**: System and Software engineering - Software Quality Requirements and Evaluation (SQuaRE) – Measurement Reference Model and Guide. Geneva, 2007.

_____. **ISO/IEC 25022**: System and Software engineering - Software Quality Requirements and Evaluation (SQuaRE) – Measurement of Quality in Use. Geneva, 2012a.

_____. **ISO/IEC 25023**: System and Software engineering - Software Quality Requirements and Evaluation (SQuaRE) – Measurement of System and Software Product Quality. Geneva, 2012b.

_____. **ISO/IEC 25030**: System and Software engineering - Software Quality Requirements and Evaluation (SQuaRE) – Quality Requirements. Geneva, 2012c.

_____. **ISO/IEC 25040**: System and Software engineering - System and software Quality Requirements and Evaluation (SQuaRE) – Evaluation process. Geneva, 2011.

ISO 25000: Series of Standards. Disponível em: <<http://iso25000.com/index.php/en/iso-25000-standards>>. Acesso em: 14 abr. 2014.

IVORY, M. Y.; HEARST, M. A. The state of the art in automating usability evaluation of user interfaces. **ACM Computing Surveys**, v. 33, n. 4, 2001, p. 470–516.

JAIN, A. K.; DUBES, R. C. **Algorithms for Clustering Data**. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1988.

JAIN, Anil K. Data clustering: 50 years beyond K-means. **Pattern recognition letters**, v. 31, n. 8, 2010, p. 651-666.

JANSEN, B. J., Search log analysis: What is it; what's been done; how to do it. **Library and Information Science Research**, v. 28, n. 3, 2006, p. 407-432.

KAROULIS, A.; POMBORTSIS, A. Usability and Learnability Evaluation of Web-Based ODL Programs. **Usability Evaluation of Online Learning Programs**. Hershey, PA: Information Science Publishing, 2003.

KELLAR, M.; WATTERS, C. Studying User Behaviour on the Web: Methods and Challenges. Workshop on Usage Analysis: Combining Logging and Qualitative Methods. In: **PROCEEDINGS OF THE Conference on Human Factors in Computing Systems**, 2005, Portland, USA, p. 2121-2122.

KRONBAUER, A. H.; SANTOS, C. A. Um Modelo de Avaliação da Usabilidade Baseado na Captura Automática de Dados de Interação do Usuário em Ambientes Reais. In: **PROCEEDINGS OF THE 10th Brazilian Symposium on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction**. Brazilian Computer Society, 2011, Porto de Galinhas, Brasil, p. 114-123.

LENAT, D. B. The nature of heuristics. **Artificial Intelligence**, v. 19, n. 2, 1982, p. 189 – 249.

LIMA, E. C.; RESENDE, A. M. P. Valores de Referencia de Métricas de Software: Uma Revisão Sistemática da Literatura. In: **PROCEEDINGS OF THE XXI Congresso de Pós-Graduação da UFLA, 2012, Lavras. XXI Congresso de Pós-Graduação da UFLA**, 2012.

MACK, R. L.; NIELSEN, J. Usability inspection methods. In: **PROCEEDINGS OF THE Conference Companion on Human Factors in Computing Systems**, 1994, New York, USA, p. 413-414.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: **PROCEEDINGS OF THE Fifth Berkeley Symposium on Mathematical Statistics and Probability**, 1967, Berkeley, California, USA.: University of California Press, 1967. v. 1, p. 281–297.

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO. **Padrões Web em Governo Eletrônico**: Cartilha de Usabilidade. Brasília, DF, 2010.

MIRKIN, B. **Clustering: a data recovery approach**. 2 ed. Boca Raton: CRC Press, 2012.

MOLICH, R.; THOMSEN, A. D.; KARYUKINA, B.; SCHMIDT, L.; EDE, M.; VAN OEL, W.; ARCURI, M. Comparative evaluation of usability tests. In: **PROCEEDINGS OF THE Conference on Human Factors in Computing Systems**, 1999, Pittsburgh, New York, p. 83–86.

MORAIS, R. M.; COSTA, A. L.; GOES, W. M.; SOMERA, S. C. Aplicação de Métricas de Software em um Modelo de Avaliação de Qualidade para Sistemas de Informação de Saúde: Um Estudo de Factibilidade em um Hospital Público Universitário. In: **ANAIS DO XVI SEMEAD - Seminários em Administração**, 2013, São Paulo.

NIELSEN, J. **Usability Engineering**. San Francisco: Morgan Kaufmann. 1994.

NIELSEN, J.; LORANGER, H. **Prioritizing Web Usability**. New York: Elsevier, New Riders Publishing. 2006.

NOY, N. F.; SHAH, N. H.; WHETZEL, P. L.; DAI, B., DORF, M., GRIFFITH, N. **BioPortal: ontologies and integrated data resources at the click of a mouse**. **Nucleic acids research**, 2009, p. 440.

OLINER, A.; GANAPATHI, A.; XU, W. Advances and challenges in log analysis. **Communications of the ACM**, v. 55, n. 2, 2012, p. 55-61.

PAGANELLI, L.; PATERNÒ, F. Intelligent analysis of user interactions with web applications. In: **PROCEEDINGS OF THE 7th International Conference on Intelligent User Interfaces**, 2002, p. 111-118.

PELLEG, D.; MOORE A. X-means: Extending K-means with Efficient Estimation of the Number of Clusters. In: **PROCEEDINGS OF THE International Conference of Machine Learning**, 2000, p. 727-734.

PLAZA, B. Google Analytics for measuring website performance. **Tourism Management**, v. 32, n. 3, 2011, p. 477-481.

PREECE, J.; ROGERS, Y.; SHARP, H. **Interaction Design: Beyond Human-Computer Interaction**. 3 ed. Chichester: Wiley, 2012.

RIVERO, L.; CONTE, T. Inspeção de usabilidade em aplicações web guiada por funcionalidades. In: **PROCEEDINGS OF THE 10th Brazilian Symposium on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction**, 2011, p. 4-5.

RIVOLLI, A.; PEREIRA, D. F. ; PANSANATO, L. T. . Capturing User Interaction with Rich Internet Applications for Usability Evaluation. In: **PROCEEDINGS OF THE 13th International Conference WWW/Internet**, 2014, Porto, p. 3-10.

RUSU, C.; RONCAGLIOLO, S.; RUSU, V.; COZZAZOS, C. A methodology to establish usability heuristics. In: **PROCEEDINGS OF THE Fourth International Conference on Advances in Computer-Human Interactions**, 2011, p. 59-62.

SANTANA, V. F.; BARANAUSKAS, M. C. C.. WELFIT: A remote evaluation tool for identifying Web usage patterns through client-side logging. **International Journal of Human-Computer Studies**, v. 76, 2015, p. 40-49.

SCHWERZ, A. L.; MORANDINI, M.; DA SILVA, S. R. A task model proposal for web sites usability evaluation for the ErgoMonitor environment. In: **PROCEEDINGS OF THE**

Human-Computer Interaction. Interaction Design and Usability, 2007. Springer Berlin Heidelberg, 2007, p. 1188–1197.

SHAH, I. Event Patterns as Indicators of Usability Problems. *Journal of King Saud University-Computer and Information Sciences*, v. 20, 2008, p. 31-43.

SHARMA, K.; SHRIVASTAVA, G.; KUMAR, V. Web mining: Today and tomorrow. In: **PROCEEDINGS OF THE Electronics Computer Technology**, 2011. p. 399-403.

SIOCHI, A. C., EHRICH, R. W. Computer Analysis of User Interfaces Based on Repetition in Transcripts of User Sessions. **ACM Transactions on Information Systems**, Vol. 9, n. 4, 1991, p. 309-335.

SWALLOW, J., HAMELUCK, D., CAREY, T. User Interface Instrumentation for Usability Analysis: A Case Study. In: **PROCEEDINGS OF THE International Conference of the Centre for Advanced Studies on Collaborative Research**, 1997, Toronto.

TIEDTKE, T.; MARTIN, C.; GERTH, N. AWUSA - A Tool for Automated Website Usability Analysis. In: **PREPROCEEDINGS OF THE 9th International Workshop on Interactive Design, Specification and Verification**, 2002, Rostock, Germany, p. 12-14.

TING, I., KIMBLE, C., KUDENCO D., UBB Mining: Finding Unexpected Browsing Behaviour in Clickstream Data to Improve a WebSite's Design. In: **PROCEEDINGS OF THE International Conference on Web Intelligence**, 2005, France, p. 179-185.

TURNBULL, D. Methodologies for Understanding Web Use with Logging in Context. Workshop on Logging Traces of Web Activity: The Mechanics of Data Collection. In: **PROCEEDINGS OF THE International Conference on World Wide Web**, 2006, Edinburgh, Scotland.

VARGAS, A.; WEFERS, H.; DA ROCHA, H. V. A method for remote and semi-automatic usability evaluation of web-based applications through users behavior analysis. In: **PROCEEDINGS OF THE 7th International Conference on Methods and Techniques in Behavioral Research**, 2010, p. 19.

VASCONCELOS, L. G.; JÚNIOR, L. A. B. USABILICS: avaliação remota de usabilidade e métricas baseadas na análise de tarefas. In: **PROCEEDINGS OF THE 10th Brazilian Symposium on on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction**, 2011, p. 303-311.

WALJI, M. F., KALENDERIAN, E., TRAN, D.; KOOKAL, K. K.; NGUYEN, V.; TOKEDE, O., PATEL, V. L., et. al. Detection and characterization of usability problems in structured data entry interfaces in dentistry. **International journal of medical informatics**, v. 82, n. 2, 2013, p. 128-138.

APÊNDICE A - Tarefas Propostas para Avaliação Automática - DREDS

Os quadros seguintes apresentam as métricas para as subcaracterísticas de apreensibilidade (Quadro 1), operabilidade e eficiência (Quadro 2), proteção ao erro do usuário (Quadro 3) e estética da interface (Quadro 4). O item “Métrica” especifica o nome atribuído à métrica relacionada à subcaracterística. O item “Descrição” apresenta o significado da métrica. O item “Heurísticas Relacionadas” descreve as heurísticas consideradas no conceito da métrica. Os itens “Procedimento” e “Pseudocódigo” especificam como a métrica é calculada. O item “Indicador” descreve como o resultado da métrica deve ser interpretado. O item “Referência” apresenta as referências utilizadas para definição da métrica.

Métrica	M ₁ - Apreensibilidade (qualidade do produto)
Descrição	Refere-se ao quanto o sistema pode ser usado por usuários específicos, para alcançar objetivos específicos com eficácia, eficiência, satisfação e sem riscos num contexto específico de uso. Embora o conceito de apreensibilidade seja semelhante ao conceito de usabilidade, o primeiro diz respeito à forma com que o usuário conclui a tarefa na primeira vez que acessa o sistema ou quando tem pouca experiência com a interface.
Heurísticas relacionadas	Como o conceito de apreensibilidade é amplo, várias heurísticas estão relacionadas a essa métrica. A primeira delas é a correlação entre o sistema e o mundo real, que define que o sistema deve evitar termos técnicos, utilizando a linguagem dos usuários e ser construído em ordem lógica e coerente. A segunda heurística está ligada ao reconhecimento da interface ao invés da memorização; a interface deve oferecer ajuda contextual e informações capazes de orientar as ações do usuário, sem precisar memorizar informações nem buscar ajuda em outras interfaces. A flexibilidade e eficiência de uso também está relacionada a essa métrica; o sistema precisa ser fácil para usuários leigos, mas flexível para se tornar ágil a usuários avançados.

Procedimento	<p>EM: quantidade média do evento carregamento da página.</p> <p>NU: número de usuários do <i>cluster</i> que finalizaram a tarefa sem buscar ajuda em outras interfaces.</p> <p>NT: número total de usuários.</p> <p>$I_1 = (NU \leq EM) / NT \rightarrow$ (indicador 1)</p> <p>NU é definido pela ocorrência do evento carregamento da página (<i>load</i>). Todas as vezes que o usuário recarrega a página ou sai do sistema para procurar ajuda em outras páginas.</p>
Indicador	Quanto mais próximo de 1 (um), maior a conformidade do sistema em relação à subcaracterística de apreensibilidade.
Pseudocódigo	<pre> início <codigo apreensibilidade > var n, nu, nt, load, cont, tempo: inteiro em, i₁: real n ← 0; nu ← 0; nt ← 0; load ← 0; cont ← 0; tempo ← 0; em ← 0; i ← 0; ler (nt) //quantidade de usuários que concluíram a tarefa enquanto tempo ≠ 9000000000 faça nt ← nt + 1 ler (nt) fim enquanto ler (em) // quantidade média do evento carregamento da página para a tarefa enquanto load ≠ 0 faça em ← em + load cont ← cont + 1 ler (em) fim enquanto em ← em / cont ler (nu) // quantidade de usuários que finalizaram a tarefa utilizando n° ≤ em enquanto n ≤ em faça nu ← nu + n ler (nu) fim enquanto para tarefa 1 até 5 faça para cluster 0 até 5 faça i₁ ← nu / nt fim <codigo apreensibilidade > </pre>
Referência	<p>(KRONBAUER; SANTOS, 2011)</p> <p>(MORAIS et al., 2013)</p> <p>(INTERNATIONAL..., 2012b)</p>

Quadro 1 – Métrica da Apreensibilidade (M₁)

Fonte: Autoria própria.

Métrica	M ₂ - Operabilidade/ Eficiência (qualidade do produto e qualidade em uso)
Descrição	Refere-se ao quanto o sistema tem atributos que o façam fácil de aprender, utilizar e controlar.

Heurísticas relacionadas	Está ligada à flexibilidade, ao projeto estético e minimalista e ao controle e liberdade do usuário. O sistema deve apresentar interface simples, clara e contextualizada. Deve ser fácil para usuários leigos e flexível e ágil para usuários avançados.
Procedimento	<p>TM: tempo médio de conclusão da tarefa proposta, sem a ocorrência de erros ou acesso a outras interfaces.</p> <p>NU: número de usuários do <i>cluster</i> que concluíram a tarefa utilizando tempo menor ou igual à média geral da tarefa.</p> <p>NT: número total de usuários.</p> <p>$I_2 = (NU \leq TM) / NT \rightarrow$ (indicador 2)</p>
Indicador	Quanto mais próximo de 1 (um), maior a conformidade do sistema em relação à subcaracterística da operabilidade/ eficiência.
Pseudocódigo	<pre> início <codigo operabilidade_eficiencia > var n, nu, nt, load, cont, tempo, e: inteiro em, tm, i₂: real n ← 0; nu ← 0; nt ← 0; load ← 0; cont ← 0; tempo ← 0; e ← 0; em ← 0; tm ← 0; i₂ ← 0; ler (nt) //quantidade de usuários que concluíram a atividade enquanto tempo ≠ 90000000000 faça nt ← nt + 1 ler (nt) fim enquanto ler (em) // quantidade média do evento carregamento da página para a tarefa enquanto load ≠ 0 faça em ← em + load cont ← cont + 1 ler (em) fim enquanto em ← (em / cont) ler (tm) // tempo médio necessário para conclusão da tarefa enquanto tempo ≠ 90000000000 e (e ≤ em) faça tm ← tm + tempo cont ← cont + 1 ler (tm) fim enquanto tm ← tm / cont ler (nu) // quantidade de usuários que finalizaram a tarefa utilizando tempo ≤ tm enquanto tempo ≤ tm faça nu ← nu + 1 ler (nu) fim enquanto para tarefa 1 até 5 faça para cluster 0 até 5 faça i₂ ← nu / nt fim <codigo operabilidade_eficiencia > </pre>
Referência	(MORAIS et al., 2013)

(INTERNATIONAL..., 2012b)

Quadro 2 – Métrica da Operabilidade/ Eficiência (M₂)**Fonte: Autoria própria.**

Métrica	M ₃ - Proteção ao erro do usuário (qualidade do produto e qualidade em uso)
Descrição	Refere-se às características do sistema para evitar erros cometidos pela interação do usuário com sistema. Dentre essas características consideram-se notas explicativas, validação de dados de entrada ou qualquer outro mecanismo para evitar que os usuários repitam ações para conclusão de uma tarefa.
Heurísticas relacionadas	Está ligada à prevenção de erros e a visibilidade do estado atual. O sistema deve ser projetado cuidadosamente, a fim de evitar que os usuários recebam mensagens de erros e sejam obrigados a repetir ações para concluir uma tarefa. Além disso, o usuário precisa saber continuamente o que está acontecendo com o sistema dentro de um tempo razoável para orientá-lo (<i>feedback</i>).
Procedimento	EM: quantidade média de eventos utilizados para conclusão de uma tarefa sem a ocorrência de erros ou acesso a outras interfaces. NU: número de usuários do <i>cluster</i> que repetiram alguma operação do sistema para finalizar a tarefa proposta, indicado pela quantidade de eventos ocorridos. NT: número total de usuários. $I_3 = 1 - ((NU \geq EM) / NT) \rightarrow (\text{indicador } 3)$
Indicador	Quanto mais próximo de 1(um), maior a adequação do sistema em relação ao impedimento de erros e à clareza dos dados.
Pseudocódigo	<pre> início <codigo protecao > var nu, nt, cont, e: inteiro em, i₂: real nu ← 0; nt ← 0; cont ← 0; e ← 0; em ← 0; i₃ ← 0; ler (nt) //quantidade de usuários que concluíram a tarefa enquanto tempo ≠ 90000000000 faça nt ← nt + 1 ler (nt) fim enquanto ler (em) // quantidade média de eventos utilizados para conclusão da tarefa enquanto e ≠ 0 faça em ← em + e cont ← cont + 1 ler (em) fim enquanto em ← (em / cont) ler (nu) // quantidade de usuários que repetiram operações para finalizar a tarefa (e ≥ em) enquanto e ≥ em faça nu ← nu + 1 ler (nu) fim enquanto </pre>

	para tarefa 1 até 5 faça para cluster 0 até 5 faça $i_3 \leftarrow 1 - (nu / nt)$ fim <codigo operabilidade_eficiencia >
Referência	(MORAIS et al., 2013) (INTERNATIONAL..., 2012b)

Quadro 3 – Métrica da Proteção ao Erro do Usuário (M₃)

Fonte: Autoria própria.

Métrica	M ₄ - Estética da interface (qualidade do produto)
Descrição	Refere-se às características do projeto de interface para garantir que o usuário tenha prazer e fique satisfeito ao interagir com o sistema. Nesse contexto, foi considerada a facilidade do usuário em encontrar a informação desejada para conclusão da tarefa e flexibilidade na execução das tarefas.
Heurística relacionada	Está ligada ao projeto estético e minimalista e flexibilidade e eficiência de uso. O sistema deve ser projetado de forma simples e clara, evitando dados irrelevantes ou desnecessários que obrigue o usuário a ficar procurando pela informação desejada e deve ser flexível e eficaz evitando que usuário se sinta frustrado na execução da tarefa.
Procedimento	MM: quantidade média da ocorrência do evento movimento do <i>mouse</i> durante a tarefa. SM: quantidade média da ocorrência do evento rolagem durante a tarefa. NU _M : número de usuários que utilizaram o movimento do <i>mouse</i> mais vezes do que MM. NU _S : número de usuários que utilizaram o evento de rolagem mais vezes que a SM. NT: número total de usuários. $I_4 = 1 - (((NU_M \geq MM) + (NU_S \geq SM)) / 2) / NT \rightarrow$ (indicador 4)
Indicador	Quanto mais próximo de 1 (um), maior a conformidade do sistema em relação à estética da interface. Usuários que passam algum tempo navegando pelo sistema, apenas utilizando o mouse ou a rolagem da tela, sem clicar ou digitar dados podem estar com dificuldades em compreender a interface ou encontrar a informação desejada (SHAH et al., 2008).
Pseudocódigo	<pre> início <codigo estetica > var nus, num, nt, cont, tempo, es, em, sm, mm, scroll, move: inteiro em, tm, i₂: real nus ← 0; num ← 0; nt ← 0; cont ← 0; tempo ← 0; es ← 0; sm ← 0; scroll ← 0; move ← 0; mm ← 0; em ← 0; tm ← 0; </pre>

	<pre> i₄ ← 0; ler (nt) //quantidade de usuários que concluíram a tarefa enquanto tempo ≠ 90000000000 faça nt← nt + 1 ler (nt) fim enquanto ler (em) // quantidade média do evento rolagem da tela utilizados para conclusão da tarefa enquanto scroll ≠ 0 faça sm ← sm + scroll cont ← cont + 1 ler (sm) fim enquanto sm← (sm / cont) ler em // quantidade média do evento movimento do mouse utilizados para conclusão da tarefa enquanto move ≠ 0 faça mm ← mm + move cont ← cont + 1 ler (mm) fim enquanto mm ← (mm / cont) ler (nus) // quantidade de usuários que utilizaram a rolagem de tela mais vezes que a média enquanto es ≥ sm faça nus ← nus + 1 ler (nus) fim enquanto ler (num // quantidade de usuários que utilizaram o movimento do mouse mais vezes que a média enquanto em ≥ mm faça num ← num + 1 ler (num) fim enquanto para tarefa 1 até 5 faça para cluster 0 até 5 faça i₄ ← 1 - (((nus + num) / 2) / nt) fim <codigo estetica > </pre>
Referência	(SHAH et al., 2008) (INTERNATIONAL..., 2012b)

Quadro 4 – Métrica da Estética da Interface (M₄).

Fonte: Autoria própria.

APÊNDICE B - Tarefas Propostas para Avaliação Automática - DREDS

TAREFA 1

Criar a história “As coisas que mais gostamos”

<i>Linha 1</i>	Narrador:	Era um daqueles dias de Verão em que sopra uma brisa suave. As nuvens corriam umas atrás das outras pelo céu, passeando pelo mundo. Milly e Molly estavam deitadas na relva, com as cabeças juntas. Fazia calor e elas estavam pensativas. E, se a Pipa e a Beta não as tivessem encontrado, provavelmente teriam adormecido. O João e o Toni também apareceram.
<i>Linha 2</i>	Molly:	— Esta é uma das minhas coisas preferidas.
<i>Linha 3</i>	João:	— O quê?
<i>Linha 4</i>	Molly:	— Deitar-me na relva a ver as nuvens. Qual é a sua coisa preferida?
<i>Linha 5</i>	João:	— Hambúrgueres.
<i>Linha 6</i>	Molly:	— Não, João, não estou falando disso. Falo de sentimentos
<i>Linha 7</i>	João:	— Como quando o meu pai vai me ver jogar voley?
<i>Linha 8</i>	Molly:	— Sim. — Eu gosto de me sentar no sofá do meu pai, quando ele se levanta e as almofadas ainda estão quentes.
<i>Linha 9</i>	Carla:	— Eu adoro sentir o cheirinho de terra molhada.
<i>Linha 10</i>	João:	— E eu gosto de encontrar a minha mãe ao pé dos morangueiros, quando chego da escola. Ela sempre me dá os melhores.
<i>Linha 11</i>	Dani:	— E eu fico feliz quando vejo filhotes brincando!
<i>Linha 12</i>	Tony:	— Quando há trovoada, gosto de deitar na cama quentinha dos meus pais.
<i>Linha 13</i>	Molly:	— Eu gosto do cheiro do bolo de chocolate saindo do forno, quando chego da escola, à sexta-feira
<i>Linha 14</i>	Narrador:	Toni ficou de pé.
<i>Linha 15</i>	João:	— É sexta-feira, Molly! O que estamos esperando? Vou chegar primeiro!
<i>Linha 16</i>	Mãe de Molly:	— Vocês são um grupo muito especial.
<i>Linha 17</i>	Narrador:	E todos disseram que sim com a cabeça.

TAREFA 2

Criar a personagem “MILLY” e incluir os seguintes diálogos nas posições indicadas:

Abaixo da *Linha 8*:

Milly:	— E eu gosto das segundas-feiras, quando os meus lençóis cheiram a sol.
--------	---

Abaixo da *Linha 11*:

Milly:	— Eu gosto de beber chocolate quente com o meu pai, ao pé da lareira, enquanto ele me conta histórias.
--------	--

Abaixo da *Linha 15*:

Milly:	— Gosto quando diz isso. Diz Milly com a boca cheia de chocolate.
--------	---

TAREFA 3

Substituir a personagem “TONY” por uma personagem nova chamada “CAROL”, com o seguinte diálogo:

Carol:	— Eu adoro sentir o abraço apertado do meu irmão mais novo.
--------	---

Substituir a personagem “MÃE DE MOLLY” por uma personagem nova chamada “DONA JULIA”, com o seguinte diálogo:

Dona Julia:	— Vocês serão amigos a vida inteira.
-------------	--------------------------------------

Observação: Mantenha os novos diálogos na mesma posição dos anteriores.

TAREFA 4

Exclua a personagem “DANI”

Exclua a personagem “CARLA”

TAREFA 5

Coloque em negrito todas as falas do “NARRADOR”

Altere para a fonte “Courier New” todas as falas da “MOLLY”

Coloque em itálico a palavra “Voley” (*Linha 7*)

APÊNDICE C - Tarefa Proposta para Avaliação Automática - SEAC

Tarefa: Transmissão de Manuscrito.

Acessar o Sistema Eletrônico de Administração de Conferências, localizado no seguinte endereço: eventos.utfpr.edu.br.

Realizar seu cadastro.

Seguir as instruções do Sistema e enviar um arquivo qualquer do seu computador. (Você pode escolher o arquivo que deseja enviar).

APÊNDICE D - Interfaces do SEAC

Figura 1 – Acesso ao cadastro do Sistema
Fonte: Sistema Eletrônico de Administração de Conferências.

Figura 2 – Página de cadastro do Sistema SEAC
Fonte: Sistema Eletrônico de Administração de Conferências.

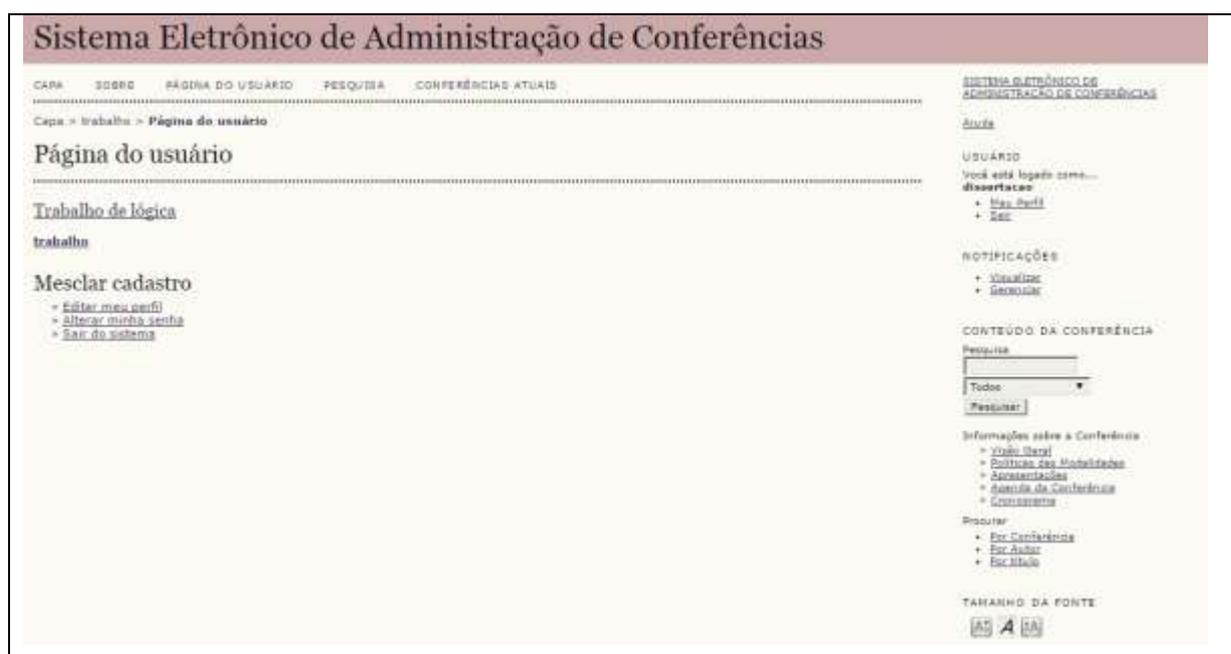


Figura 3 – Página para escolha da conferência
Fonte: Sistema Eletrônico de Administração de Conferências.

Na Figura 3, o usuário que já está cadastrado e conectado ao sistema pode escolher qual conferência que participar. Nesse caso, as duas conferências disponíveis são “Trabalho de Lógica” e “trabalho”.



Figura 4 – Página com as opções para a conferência “trabalho”
Fonte: Sistema Eletrônico de Administração de Conferências.

Na interface apresentada na Figura 4, o usuário precisa escolher a opção “Chamada para Submissões”.



Figura 5 – Página de início do processo de submissão de arquivo
Fonte: Sistema Eletrônico de Administração de Conferências.

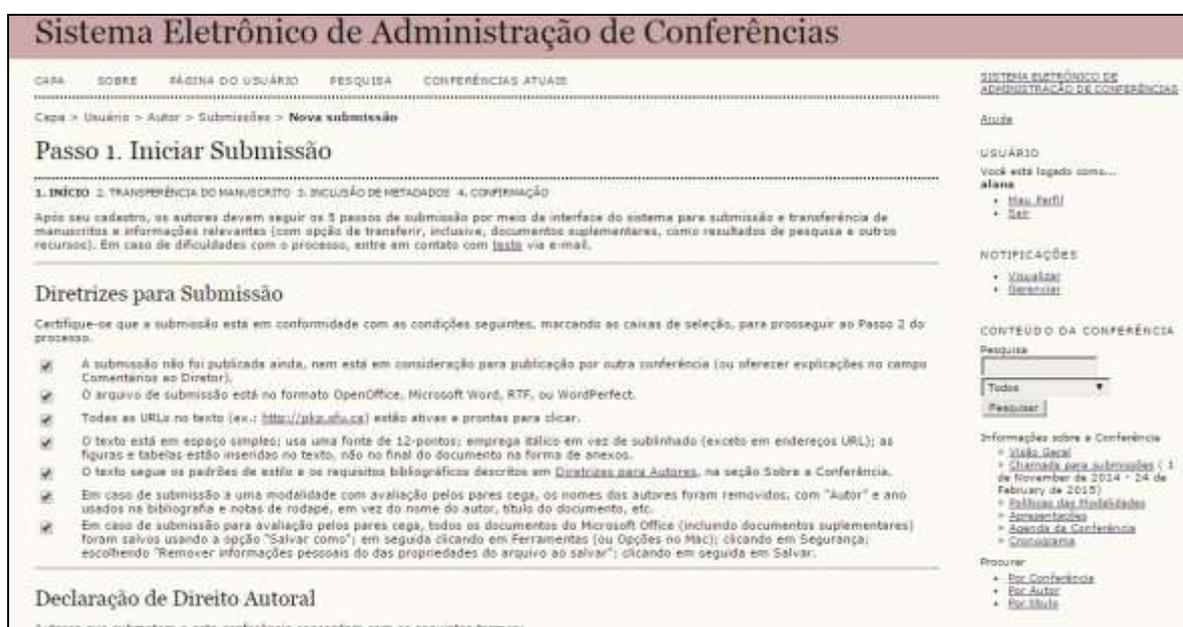


Figura 6– Página do processo de submissão de arquivo
Fonte: Sistema Eletrônico de Administração de Conferências.

Sistema Eletrônico de Administração de Conferências

CAPA | SOBRE | PÁGINA DO USUÁRIO | PESQUISA | CONFERÊNCIAS ATUAIS

Capa > Usuário > Autor > Submissões > Nova submissão

Passo 2. Transferência do Manuscrito

1. INÍCIO 2. TRANSFERÊNCIA DO MANUSCRITO 3. INCLUSÃO DE METADADOS 4. CONFIRMAÇÃO

Para transferir um manuscrito para o sistema, execute os seguintes passos:

1. Nesta página, clique em Procurar (Browse, se o idioma do navegador for inglês, ou Escolher Arquivo) para abrir uma janela de seleção. Escolha Arquivo (Enviar Arquivo) para localizar o documento no disco rígido do seu computador.
2. Localize o documento desejado e selecione-o.
3. Clique em Abrir na janela, que colocará o nome do documento no campo respectivo do formulário de envio.
4. Clique em Transferir, para enviar o documento do seu computador para o servidor de hospedagem do sistema. O sistema dará um novo nome ao documento seguindo os padrões pré-definidos.
5. Uma vez transferido, clique em Salvar e Continuar no final da página.

Após seu cadastro, os autores devem seguir os 5 passos de submissão por meio da interface do sistema para submissão e transferência de manuscritos e informações relevantes (com opção de transferir, inclusive, documentos suplementares, como resultados de pesquisa e outros recursos). Em caso de dificuldades com o processo, entre em contato com [igata](mailto:igata@ufpa.br) via e-mail.

Documento da Submissão

Nenhum documento transferido.

Transferir documento da submissão Nenhum arquivo selecionado

SISTEMA ELETRÔNICO DE ADMINISTRAÇÃO DE CONFERÊNCIAS

Ajuda

USUÁRIO
Você está logado como...
alana

- Meu Perfil
- Sair

NOTIFICAÇÕES

- Visualizar
- Gerenciar

CONTEÚDO DA CONFERÊNCIA

Busque

Todos

Informações sobre a Conferência

- Visão Geral
- Chamada para submissões (1 de November de 2014 - 24 de February de 2015)
- Políticas das Modalidades
- Apresentação
- Sobre a Conferência
- Sistema

Procurar

- Por Conferência
- Por Autor
- Por Título

Figura 7 – Página para escolha do arquivo a ser submetido
Fonte: Sistema Eletrônico de Administração de Conferências.

Sistema Eletrônico de Administração de Conferências

CAPA | SOBRE | PÁGINA DO USUÁRIO | PESQUISA | CONFERÊNCIAS ATUAIS

Capa > Usuário > Autor > Submissões > Nova submissão

Passo 2. Transferência do Manuscrito

1. INÍCIO 2. TRANSFERÊNCIA DO MANUSCRITO 3. INCLUSÃO DE METADADOS 4. CONFIRMAÇÃO

Para transferir um manuscrito para o sistema, execute os seguintes passos:

1. Nesta página, clique em Procurar (Browse, se o idioma do navegador for inglês, ou Escolher Arquivo), abrindo uma janela de seleção. Escolha Arquivo (Enviar Arquivo) para localizar o documento no disco rígido do seu computador.
2. Localize o documento desejado e selecione-o.
3. Clique em Abrir na janela, que colocará o nome do documento no campo respectivo do formulário de envio.
4. Clique em Transferir, para enviar o documento do seu computador para o servidor de hospedagem do sistema. O sistema dará um novo nome ao documento seguindo os padrões pré-definidos.
5. Uma vez transferido, clique em Salvar e Continuar no final da página.

Após seu cadastro, os autores devem seguir os 5 passos de submissão por meio da interface do sistema para submissão e transferência de manuscritos e informações relevantes (com opção de transferir, inclusive, documentos suplementares, como resultados de pesquisa e outros recursos). Em caso de dificuldades com o processo, entre em contato com [igata](mailto:igata@ufpa.br) via e-mail.

Documento da Submissão

Nome do documento	30-07-1-0H.pdf
Nome original do Documento	Calendario Acadêmico 2015.pdf
Tamanho do documento	229KB
Data de transferência	2015-02-12 10:52 AM

Substituir documento da submissão Nenhum arquivo selecionado

SISTEMA ELETRÔNICO DE ADMINISTRAÇÃO DE CONFERÊNCIAS

Ajuda

USUÁRIO
Você está logado como...
alana

- Meu Perfil
- Sair

NOTIFICAÇÕES

- Visualizar
- Gerenciar

CONTEÚDO DA CONFERÊNCIA

Busque

Todos

Informações sobre a Conferência

- Visão Geral
- Chamada para submissões (1 de November de 2014 - 24 de February de 2015)
- Políticas das Modalidades
- Apresentação
- Sobre a Conferência
- Sistema

Procurar

- Por Conferência
- Por Autor
- Por Título

Figura 8 – Página de confirmação do arquivo escolhido
Fonte: Sistema Eletrônico de Administração de Conferências.

Sistema Eletrônico de Administração de Conferências

CAPA SOBRE PÁGINA DO USUÁRIO PESQUISA CONFERÊNCIAS ATUAIS

Capa > Usuário > Autor > Submissões > Nova submissão

Passo 3. Metadados da Submissão (Indexação)

1. INÍCIO 2. TRANSFERÊNCIA DO MANUSCRITO 3. INCLUSÃO DE METADADOS 4. CONFIRMAÇÃO

Autores

Prenome*
 Nome do meio
 Sobrenome*
 E-mail*
 URL
 Instituição
 País (Sua instituição, por exemplo "Simon Fraser University")
 Resumo da Biografia (Ex.: departamento e área)

SISTEMA ELETRÔNICO DE ADMINISTRAÇÃO DE CONFERÊNCIAS

Busca

USUÁRIO
 Você está logado como...
alana
 • Meu Perfil
 • Sair

NOTIFICAÇÕES
 • Visualizar
 • Gerenciar

CONTEÚDO DA CONFERÊNCIA
 Pesquisa
 Todos

Informações sobre a Conferência
 • [Vida Geral](#)
 • [Chamada para submissões \(1 de November de 2014 - 24 de February de 2015\)](#)
 • [Políticas das Modalidades](#)
 • [Anúncios](#)
 • [Agenda da Conferência](#)
 • [Contato](#)

Procurar
 • [Por Conferência](#)
 • [Por Autor](#)
 • [Por Título](#)

Figura 9 – Página de inclusão de dados sobre a submissão
 Fonte: Sistema Eletrônico de Administração de Conferências.

Sistema Eletrônico de Administração de Conferências

CAPA SOBRE PÁGINA DO USUÁRIO PESQUISA CONFERÊNCIAS ATUAIS

Capa > Usuário > Autor > Submissões > Nova submissão

Passo 5. Confirming the Submission

1. INÍCIO 2. TRANSFERÊNCIA DO MANUSCRITO 3. INCLUSÃO DE METADADOS 4. CONFIRMAÇÃO

Após concluídos e verificados os passos anteriores, clique em Concluir Submissão para enviar seu manuscrito para a conferência Trabalho de lógica. Um e-mail de confirmação será enviado e a situação da submissão, dentro do processo editorial, pode ser acompanhada entrando no sistema. Agradecemos seu interesse em contribuir com trabalhos para a conferência Trabalho de lógica.

Resumo de Documentos

ID	NOME ORIGINAL DO DOCUMENTO	TIPO	TAMANHO DO DOCUMENTO	DATA DE TRANSFERÊNCIA
87	CALENDARIO ACADEMICO 2015/2016	Documento de Submissão	228KB	02-12

SISTEMA ELETRÔNICO DE ADMINISTRAÇÃO DE CONFERÊNCIAS

Busca

USUÁRIO
 Você está logado como...
alana
 • Meu Perfil
 • Sair

NOTIFICAÇÕES
 • Visualizar
 • Gerenciar

CONTEÚDO DA CONFERÊNCIA
 Pesquisa
 Todos

Informações sobre a Conferência
 • [Vida Geral](#)
 • [Chamada para submissões \(1 de November de 2014 - 24 de February de 2015\)](#)
 • [Políticas das Modalidades](#)
 • [Anúncios](#)
 • [Agenda da Conferência](#)
 • [Contato](#)

Procurar
 • [Por Conferência](#)
 • [Por Autor](#)
 • [Por Título](#)

Figura 10 – Página de confirmação da submissão
 Fonte: Sistema Eletrônico de Administração de Conferências.

Sistema Eletrônico de Administração de Conferências

CAPA SOBRE PÁGINA DO USUÁRIO PESQUISA CONFERÊNCIAS ATIVAS

Capa > Usuário > Autor > Submissões > Submissões Ativas

Submissões Ativas

Submissão concluída. Agradecemos seu interesse em contribuir com manuscritos para a conferência Trabalho de lógica.

[- Submissões Ativas](#)

SISTEMA ELETRÔNICO DE ADMINISTRAÇÃO DE CONFERÊNCIAS

AJUDA

USUÁRIO
 Você está logado como...
 alana
 • Meu perfil
 • Sair

NOTIFICAÇÕES
 • Visualizar
 • Remover

CONTEÚDO DA CONFERÊNCIA
 Pesquisa

 Todas ▾

Informações sobre a Conferência
 • Visão Geral
 • Chamada para submissões (1 de novembro de 2014 - 24 de February de 2015)
 • Política das finalidades
 • Submissões
 • Agenda da Conferência
 • Cronograma

Buscar:
 • Por Conferência
 • Por Autor
 • Por título

Figura 11 – Página de conclusão da submissão
Fonte: Sistema Eletrônico de Administração de Conferências.

APÊNDICE E - Resultado da Avaliação Heurística - DREDS

Descrição do Problema de Usabilidade	Nível de Severidade
Heurística 1 - Visibilidade do estado atual do sistema	
Ausência de informação sobre qual formatação está aplicada ao texto selecionado.	2
Ao aplicar novas formatações, os botões de formatação anteriores continuam habilitados.	3
Botão “ <i>remove formatting</i> ” não remove toda a formatação do texto.	2
Ausência de destaque para a ação primária proeminente.	2
Heurística 2 - Correlação entre o sistema e o mundo real	
Funcionalidade descrita de forma incorreta. A opção “ <i>add new scene</i> ” adiciona um novo diálogo e não uma nova cena.	2
O sistema possui palavras em inglês.	1
Heurística 3 – Controle e liberdade do usuário	
Impossibilidade de remover a formatação selecionada (opções de sobrescrito, subscrito, cor de fundo).	2
Ao aplicar novas formatações, os botões de formatação anteriores continuam habilitados.	3
Impossibilidade de excluir uma personagem e seus diálogos de uma só vez. O sistema exige a exclusão de um diálogo de cada vez.	2
Possibilidade de iniciar uma nova história sem sair da página.	3
Botão “ <i>remove formatting</i> ” não remove toda a formatação do texto.	2
Ausência do botão “desfazer alterações”.	3
Impossibilidade de exportar/importar conteúdo.	3
Ausência de botão/tela de confirmação de exclusão de cena.	3
Ausência da opção de “salvar” a história.	2
Ausência do botão de “salvar” para alterações nos diálogos. Basta clicar em qualquer lugar fora da caixa de texto.	2
Possibilidade de alterar o tipo de fonte mesmo sem qualquer texto selecionado.	1
Heurística 4 – Consistência e padrões	
Ausência de informação sobre qual formatação está aplicada ao texto selecionado.	2
Ao aplicar novas formatações, os botões de formatação anteriores continuam habilitados.	3
Ausência do botão de “salvar” para alterações nos diálogos. Basta clicar em qualquer lugar fora da caixa de texto.	2
Ausência da opção de confirmar exclusões de diálogos.	3
A opção de excluir uma personagem somente é visível ao passar o mouse sobre a opção.	1
Ausência de botão/tela de confirmação de exclusão de cena.	3
Ausência da opção de “salvar” a história.	2

Heurística 5 – Prevenção de erros	
Impossibilidade de remover a formatação selecionada (opções de sobrescrito, subscrito, cor de fundo).	2
Possibilidade de incluir novos diálogos sem a identificação da personagem correspondente.	3
Impossibilidade de formatação do texto do título.	2
Possibilidade de iniciar uma nova história sem sair da página.	3
Possibilidade de criar um diálogo sem texto (linhas em branco).	3
Botão “ <i>remove formatting</i> ” não remove toda a formatação do texto.	2
Ausência do botão “desfazer alterações”.	3
Ausência de destaque para a ação primária proeminente.	2
Ausência de botão/tela de confirmação de exclusão de cena.	3
Ausência da opção de “salvar” a história.	2
Ausência do botão de “salvar” para alterações nos diálogos. Basta clicar em qualquer lugar fora da caixa de texto.	2
Ao aplicar novas formatações, os botões de formatação anteriores continuam habilitados.	3
Ausência de informação sobre qual formatação está aplicada ao texto selecionado.	2
Heurística 6 – Reconhecimento ao invés de memorização	
A opção de excluir uma personagem somente é visível ao passar o mouse sobre a opção.	1
Ausência de informações sobre como criar diálogos e alterar personagens.	2
Heurística 7 – Flexibilidade e eficiência de uso	
Impossibilidade de formatação do texto do título.	2
Ausência de destaque para a ação primária proeminente.	2
Impossibilidade de exportar/importar conteúdo.	3
Necessidade de definir a formatação a cada diálogo. A formatação anterior não fica salva.	2
Impossibilidade de formatação do texto do título.	2
Botão “ <i>remove formatting</i> ” não remove toda a formatação do texto.	2
Ausência do botão “desfazer alterações”.	3
Ausência da opção de “salvar” a história.	2
Ausência do botão de “salvar” para alterações nos diálogos. Basta clicar em qualquer lugar fora da caixa de texto.	2
Heurística 8 – Projeto estético e minimalista	
Ausência de limite lateral para textos longos. Não existe quebra de linha automática.	1
Ausência do botão “desfazer alterações”.	3
Ausência de botão/tela de confirmação de exclusão de cena.	3
Heurística 8 – Projeto estético e minimalista	
Ausência da opção de “salvar” a história.	2
Ausência do botão de “salvar” para alterações nos diálogos. Basta clicar em qualquer lugar fora da caixa de texto.	2

Ausência da opção de confirmar exclusões de diálogos.	3
Impossibilidade de formatação do texto do título.	2
Heurística 9 – Suporte aos usuários no reconhecimento, diagnóstico e recuperação de erros	
Nenhum dos avaliadores apontou problemas de usabilidade relacionados a esta heurística.	
Heurística 10 – Informações de ajuda e documentação	
Ausência de descrição da funcionalidade da aplicação.	1
Ausência de páginas de ajuda/documentação.	2
Ausência de página de contato com o desenvolvedor.	1
Ausência de informações sobre como criar diálogos e alterar personagens.	2

Quadro 1 – Resultado da Avaliação Heurística do DREDS

Fonte: Autoria própria.

APÊNDICE F - Resultado da Avaliação Heurística - SEAC

Descrição do Problema de Usabilidade	Nível de Severidade
Heurística 1 - Visibilidade do estado atual do sistema	
O acesso ao sistema fora da rede onde está hospedado, eleva o tempo de resposta ao usuário.	2
Quando o sistema é carregado, automaticamente aparece uma página com o título “trabalho”, no entanto não é claro o que significa esse trabalho.	3
Não é claro porque ao carregar a página inicial do sistema, o menu de localização aparece com dois níveis de acesso (Capa > Trabalho de lógica > trabalho).	3
Os links do menu principal (Capa, Sobre, Acesso, ...) podem fazer ou não fazer parte da hierarquia de navegação relacionada ao tópico “trabalho” (o evento inicial), quando deveriam ser links de segundo nível de profundidade.	3
Na tela inicial não fica claro sobre o que é o sistema, qual sua finalidade e em que ponto o usuário está.	4
Após realizado o cadastro, a tela é muito confusa e não fica claro o que o usuário deve realizar.	4
Após realizar o cadastro não é apresentada uma mensagem de que tudo ocorreu bem.	2
Heurística 1 - Visibilidade do estado atual do sistema	
Após concluir a submissão, a mensagem de sucesso informada é muito apagada, poderia ser melhor apresentada.	1
Heurística 2 - Correlação entre o sistema e o mundo real	
Na tela inicial, algumas informações estão em outro idioma, como as datas.	4
O formulário de cadastro possui um campo chamado “prenome” e isso pode não fazer muito sentido para o usuário.	4
O formulário de cadastro possui no campo sexo as iniciais M e F. As palavras poderiam ser escritas por extenso.	2
Não fica claro como a funcionalidade de pesquisa funciona, onde a busca é realizada.	3
O sistema sempre retorna que nenhum resultado foi encontrado na funcionalidade de pesquisa.	3
Heurística 3 – Controle e liberdade do usuário	
Não há opção de confirmação da ação, ao clicar no botão “cancelar”.	4
Não há opção de confirmação da ação, ao clicar no botão “concluir a submissão”.	4
Heurística 4 – Consistência e padrões	
Em uma tela o sistema mostra o <i>link</i> “clique aqui para iniciar o processo de submissão” e em outra tela, um link com a mesma função é chamado de “nova submissão”.	2
Na mesma tela existe a mesma informação escrita de duas formas: “Passo 5. Confirming the Submission” e “4. Confirmação”.	3

O <i>link</i> de mapa do portal fica localizado dentro da página de informações sobre a conferência. No entanto, o sistema poderia gerenciar múltiplas conferências.	3
Heurística 5 – Prevenção de erros	
No formulário de cadastro, a opção “cadastrar como” não deixa claro qual opção o usuário deve selecionar.	3
No formulário de cadastro, embora o botão “criar” tenha um destaque de cor ele é menor do que o botão “cancelar”.	2
Na tela de cadastro os campos não realizam restrições, como limitar os caracteres aceitos no campo de <i>login</i> ou tamanho do campo senha.	4
Todos os <i>links</i> são apresentados da mesma forma, contudo, alguns links poderiam receber um destaque devido a sua importância, por exemplo, o <i>link</i> “nova submissão”.	4
Heurística 6 – Reconhecimento ao invés de memorização	
Na tela inicial, o botão “cadastrar” está em uma lista de opções sem uma sequência lógica de ações e não está próximo do formulário de <i>login</i> .	3
Na tela inicial de submissão, a marcação das etapas (etapa atual e etapas a serem seguidas) poderia receber um destaque melhor.	1
Heurística 7 – Flexibilidade e eficiência de uso	
Na funcionalidade de pesquisa, quando os resultados aparecem, não existe a opção para a realização de uma nova busca.	3
O sistema espera uma submissão completa de dados para fornecer mensagens de erro. Essa validação poderia ser automatizada no lado cliente da aplicação.	3
Heurística 8 – Projeto estético e minimalista	
O formulário de cadastro possui muitos campos que não são obrigatórios e poderiam estar em uma segunda fase, opcional.	3
A tela que contém o formulário de cadastro possui muitas informações extras nas barras laterais.	2
Na tela de submissão, etapa 1, há muita informação para o usuário, deixando a interface poluída.	3
A funcionalidade de pesquisa poderia ter reduzido o número de possibilidades de busca. São muitos campos que podem confundir o usuário. Poderia ser elaborada uma funcionalidade de busca avançada, mas em uma página separada.	3
Heurística 9 – Suporte aos usuários no reconhecimento, diagnóstico e recuperação de erros	
Ao submeter o formulário com erro, a mensagem de erro é exibida longe do campo em que o erro ocorreu.	3
Na tela de submissão, etapa 1, ao tentar avançar sem marcar as diretrizes, uma mensagem de erro é informada. No entanto, a mensagem não deixa claro que os erros referem-se às diretrizes.	3
Heurística 10 – Informações de ajuda e documentação	
A tela de ajuda está em outro idioma e não possui instruções claras.	4
Na tela inicial, os <i>links</i> referentes a informações sobre a conferência não possuem informações relevantes.	3
Na tela de cadastro e na tela de submissão, a informação de que campos marcados com * são obrigatórios deveria estar no início e não no final.	2

Na tela de submissão, em todas as etapas, a ajuda é apresentada no mesmo formato (cor e tamanho) do texto relacionado aos elementos do formulário de envio, deixando a informação confusa.	3
O mapa do portal parece estar inconsistente com relação aos níveis de navegação possíveis.	2

Quadro 1 – Resultado da Avaliação Heurística do SEAC**Fonte: Autoria própria.**