

Anabela Prista Saraiva Serrano

Sistemas de Informação e Qualidade:
a avaliação de arquivos e bibliotecas
digitais

Tese submetida à Faculdade de Letras da Universidade do Porto para obtenção do grau de Doutor em Informação e Comunicação em Plataformas Digitais conferido pela Universidade do Porto e pela Universidade de Aveiro

Orientadora: Prof. Doutora Cândida Fernanda Antunes Ribeiro
(Faculdade de Letras da Universidade do Porto)

Coorientador: Prof. Doutor Joaquim Manuel Henriques de Sousa Pinto
(Universidade de Aveiro)



Porto
2013

Júri das Provas de Doutoramento em Informação e Comunicação em Plataformas Digitais, requeridas por Anabela Prista Saraiva Serrano, nomeado pelos despachos Reitorais da Universidade do Porto de 17 de Fevereiro de 2014 e de 30 de Setembro de 2014:

Presidente:

Doutora Maria da Graça Lisboa Castro Pinto, Professora Catedrática do Departamento de Estudos Portugueses e Estudos Românicos da Faculdade de Letras da Universidade do Porto (por delegação reitoral).

Vogais:

Doutora Maria Manuel Lopes Figueiredo Costa Marques Borges, Professora Auxiliar do Departamento de Filosofia, Comunicação e Informação da Faculdade de Letras da Universidade de Coimbra;

Doutor Pedro Manuel Rangel Santos Henriques, Professor Associado da Escola de Engenharia da Universidade do Minho;

Doutor Óscar Emanuel Chaves Mealha, Professor Associado com Agregação do Departamento de Comunicação e Arte da Universidade de Aveiro;

Doutor Joaquim Manuel Henriques de Sousa Pinto, Professor Auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro;

Doutora Cândida Fernanda Antunes Ribeiro, Professora Catedrática do Departamento de Jornalismo e Ciências da Comunicação da Faculdade de Letras da Universidade do Porto;

Doutor Armando Manuel Barreiros Malheiro da Silva, Professor Associado do Departamento de Jornalismo e Ciências da Comunicação da Faculdade de Letras da Universidade do Porto.

Projeto de investigação financiado pela FCT no âmbito do Programa PROTEC 2008 (SFRH/BD/49547/2009)

FCT Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR Portugal

Dedico esta tese:

À minha mãe, e amiga, que me educou e orientou incentivando-me a ser sempre melhor e a concretizar os meus objetivos sem temer obstáculos, sem descurar os princípios éticos e morais, e que muito contribuiu na sua atividade profissional como docente para uma sociedade em que se materialize menos a vida e em que se valorize o conhecimento e o sentimento, uma sociedade mais culta e esclarecida, mais justa e interventiva, mais transparente e leal.

Ao meu pai, e companheiro de descobertas da infância, cujos sonhos para uma menina de 6 anos, ainda, não cumpri.

Sumário

Agradecimentos	9
Resumo e palavras-chave	11
Abstract and keywords	12
Índice de figuras	13
Glossário de acrónimos e siglas	17
Introdução	19
1. Contextualização do projeto de investigação	19
2. Objeto de estudo e objetivos de investigação	25
3. Questão de investigação e hipóteses	26
4. Método de investigação	27
5. Fontes de informação	29
6. Organização da tese	29
Capítulo 1 – A avaliação de arquivos e bibliotecas digitais: o estado da arte ..	31
1.1. – A avaliação de arquivos e bibliotecas digitais – estado da arte	31
1.2. – A usabilidade em arquivos e bibliotecas digitais – estado da arte	57

Capítulo 2 – Os sistemas de informação digitais e o acesso à informação: conceitos e contextualização científica	79
2.1. – Os sistemas de informação e os repositórios digitais.....	79
2.2. – A interface do utilizador e a pesquisa por assunto.....	86
2.3. - Os sistemas de organização do conhecimento.....	89
2.4. - Os sistemas de recuperação da informação	100
2.4.1. - A pesquisa e recuperação da informação como área científica	100
2.4.2. – A pesquisa da informação	101
2.4.3. – A recuperação da informação	104
2.4.4. – O comportamento informacional e os estudos do utilizador.....	108
Capítulo 3 – O sistema QUETTA ou um Modelo Integrado de Estudo do Comportamento do Utilizador.....	111
3.1. – Os arquivos e as bibliotecas digitais em Portugal: a constituição da amostra e a sua caracterização	111
3.2. – A conceção do sistema QUETTA.....	129
3.3. – A recolha de dados gerada pelo sistema QUETTA	156
3.3.1. – Os questionários.....	163
3.3.2. – Os mapas de fixações e os <i>heatmaps</i>	166
3.3.3. – O <i>think-aloud</i>	169
Capítulo 4 – Análise dos dados e discussão dos resultados da investigação... 171	171
4.1. – Análise dos dados.....	171
4.1.1. – Questionários	172
4.1.1.1. - Caracterização demográfica do participante no estudo QUETTA.....	178
4.1.1.2. - Caracterização das competências informacionais do participante no estudo QUETTA.....	179
4.1.2. - Testes de usabilidade com <i>eyetracking</i>	184
4.1.3. - <i>Think-aloud</i>	192
4.2. – Discussão dos resultados.....	194

4.2.1. – A acessibilidade dos SOC nos SI digitais do tipo arquivo digital e do tipo biblioteca digital.....	198
4.2.2. - Os SOC como facilitadores da recuperação da informação por assunto.....	200
4.2.3. - A influência positiva do conhecimento prévio dos SOC pelo utilizador como fator de sucesso e de eficiência na pesquisa por assunto	200
4.2.4. - A influência da presença de SOC no SI digital no grau de satisfação do utilizador para com o mesmo	201
Conclusão	203
1. Considerações gerais.....	203
2. Limitações	204
3. Investigação futura.....	205
4. Contributos.....	205
Referências bibliográficas	207
Anexos	229
Anexo 1 – Tabela 1 - Projetos e programas de investigação em avaliação de serviços de informação.....	231
Anexo 2 - Tabela 2 - Critérios de avaliação da usabilidade por autor(es) e por ano.....	235
Anexo 3 – Tabela 3 - Repositórios digitais sediados em Portugal em 2009 com registo no ROAR e no OpenDOAR.....	239
Anexo 4 – Modelo do Sistema QUETTA.....	243
Anexo 5 – Atividade do <i>site</i> QUETTA.....	245
Anexo 6 – Questionários.....	247

Agradecimentos

À minha orientadora Prof. Doutora Fernanda Ribeiro pela disponibilidade inextinguível e pelos valiosos conselhos neste percurso trilhado desde há alguns anos, pelo incentivo que sempre me deu, sobretudo desde 2001, em iniciar um percurso científico, bem como pelo exemplo de perfil académico e científico de elevado impacto na Ciência da Informação que me influencia desde que nos conhecemos em 1989.

Ao meu coorientador Prof. Doutor Joaquim Sousa Pinto, igualmente, pela disponibilidade inextinguível e pelos valiosos conselhos, e pelo apoio no decurso do trabalho laboratorial incluindo as diligências que conduziram à cedência do equipamento da Portugal Telecom (PT). O seu percurso académico e profissional, nomeadamente, nas bibliotecas e arquivos digitais tem sido uma inspiração desde que nos conhecemos em 2005.

Ao Prof. Doutor Óscar Mealha e ao Prof. Doutor Fernando Ramos pelo apoio dado na disponibilização do equipamento que permitiu a concretização deste projeto de investigação e pelas suas palavras no “*Summer Doctoral Consortium*” de 2010 na Universidade de Verão da Universidade do Porto.

Ao Prof. Doutor Armando Malheiro da Silva pela sua acutilância científica e pela sua personalidade afável, generosa e marcante.

Ao CETAC.Media (Pólo de Aveiro) e à Prof. Doutora Ana Veloso pela cedência das instalações e equipamento, e pela orientação inicial na utilização do *eyetracker*, bem como à PT que cedeu o equipamento à Universidade de Aveiro no âmbito de projetos de investigação e do Sapo Labs.

Ao colega Samuel Almeida do CETAC.Media (Pólo de Aveiro), e do ICPD, pelo apoio dado nas muitas horas diárias de laboratório concentradas em duas semanas intensas de imenso trabalho. À Dra. Carla Amaral dos Serviços Académicos da Faculdade de Letras da Universidade do Porto pela sua diligência.

Aos voluntários que participaram no estudo, e aos que solicitaram agendamento ou reagendamento para a semana em que o equipamento já estaria em Lisboa por não poderem comparecer nas datas e/ou horas previstas inicialmente, um enorme agradecimento pela generosa disponibilidade e pelo rigoroso cumprimento de todas as instruções no decurso do teste de usabilidade com *eyetracking*, bem como pelas observações realizadas e pelos cumprimentos que dirigiram ao projeto. Aos colegas do IEETA, do ICPD e da Biblioteca da UA que participaram no estudo, um especial agradecimento pelo apoio e convívio que ajudaram a combater o cansaço e tornaram aquelas semanas únicas e inesquecíveis.

À Fundação para a Ciência e Tecnologia (FCT) pela concessão do financiamento para concretizar este projeto de investigação, no âmbito do Programa PROTEC 2008, à candidatura SFRH/BD/49547/2009.

À Presidência do Instituto Politécnico do Porto (IPP) por ter apoiado este projeto de investigação subscrevendo a candidatura ao PROTEC 2008 de uma docente da unidade orgânica de Vila do Conde, a Escola Superior de Estudos Industriais e de Gestão (ESEIG).

À Presidência da ESEIG por ter incentivado e apoiado as candidaturas ao Programa PROTEC 2008 da FCT.

Aos meus alunos e aos meus colegas por todas as experiências que me proporcionaram ao longo destes anos e pelas inúmeras amizades que ficam para toda a vida.

Aos meus amigos e familiares que me aconselharam e apoiaram, neste percurso, pelo carinho e pela generosidade. A todos um enorme abraço de reconhecimento e amizade.

Resumo

O projeto de investigação desenvolvido, e aqui relatado, teve por objetivo efetuar uma avaliação da qualidade dos sistemas de informação do tipo arquivo digital e biblioteca digital, em Portugal, quanto à presença de sistemas de organização do conhecimento e quanto à sua acessibilidade na interface do utilizador daquele tipo de sistemas de informação. Os arquivos e as bibliotecas digitais constituem uma realidade recente e encontram-se em fase de crescimento, necessitando de ser avaliados para que se melhore o seu desenvolvimento sob a perspetiva do utilizador. Para a concretização deste projeto, foi desenvolvido um modelo integrado de estudo do comportamento do utilizador que combina várias técnicas de investigação em ambiente laboratorial e que poderá ser aplicado a outros sistemas de informação digitais. O sistema QUETTA combina três técnicas de investigação em que a componente central pertence ao método experimental – os testes de usabilidade com *eyetracking* - e se alicerça no inquérito – os questionários - e no método de observação - o *think-aloud*. O estudo realizado com o sistema QUETTA permitiu demonstrar a importância dos sistemas de organização do conhecimento nos repositórios digitais e o seu impacto na sociedade da informação e do conhecimento, bem como a necessidade de os diversificar e de os melhorar aumentando a sua presença naqueles sistemas de informação digitais.

Palavras-chave: Sistemas de Informação; Avaliação de bibliotecas digitais; Avaliação de arquivos digitais; Sistemas de organização do conhecimento; Modelo de avaliação; Sistema QUETTA; Recuperação da informação; Testes de usabilidade com *eyetracking*; Questionários; *Think-Aloud*; Qualidade; Acesso à informação; Comportamento informacional; Indexação

Abstract

The research project developed and reported here aimed to make an assessment of the quality of the information systems of the type digital library and digital archive, in Portugal, for the presence of knowledge organization systems and for their accessibility at the user interface of that kind of information systems. The digital libraries and archives are a recent phenomenon and are in the growth phase and need to be evaluated in order to improve its development under the perspective of the user. To implement this project, an integrated model for studying user behavior that combines various techniques of research in the laboratory, and that may be applied to other digital information systems, was developed. The Quetta system combines three research techniques in which the central component belongs to the experimental method - eyetracking usability tests - and is underpinned by the survey - questionnaires - and by the method of observation - the think-aloud. The Quetta system study allowed demonstration of the importance of knowledge organization systems in digital repositories and their impact on the information and knowledge society and the need to diversify and improve them by increasing their presence in those digital information systems.

Keywords: Information Systems; Evaluation of digital libraries; Evaluation of digital archives; Knowledge organization systems; Evaluation model; Quetta system; Information retrieval; Eyetracking usability tests; Questionnaires; Think-Aloud; Quality; Information access; Information behavior; Indexing

Índice de figuras

FIGURA 1 – GRÁFICO COMPARATIVO DE DADOS DE DEZ PAÍSES (BRASIL, CHINA, ALEMANHA, HONG KONG, HOLANDA, PORTUGAL, COREIA DO SUL, SUÍÇA, REINO UNIDO E ESTADOS UNIDOS DA AMÉRICA) QUANTO À VELOCIDADE MÉDIA DE ACESSO À INTERNET NO PRIMEIRO TRIMESTRE DE 2013.	44
FIGURA 2 - GRÁFICO COMPARATIVO DE DADOS DE TRÊS PAÍSES DA UNIÃO EUROPEIA (ALEMANHA, PORTUGAL E REINO UNIDO) QUANTO À VELOCIDADE MÉDIA DE ACESSO À INTERNET NO PRIMEIRO TRIMESTRE DE 2013.	45
FIGURA 3 - GRÁFICO COMPARATIVO DE DADOS DE DEZ PAÍSES (BRASIL, CHINA, ALEMANHA, HONG KONG, HOLANDA, PORTUGAL, COREIA DO SUL, SUÍÇA, REINO UNIDO E ESTADOS UNIDOS DA AMÉRICA) QUANTO À VELOCIDADE MÉDIA DE PICOS DE ACESSO À INTERNET NO PRIMEIRO TRIMESTRE DE 2013.	46
FIGURA 4 - GRÁFICO COMPARATIVO DE DADOS DE TRÊS PAÍSES DA UNIÃO EUROPEIA (ALEMANHA, PORTUGAL E REINO UNIDO) QUANTO À VELOCIDADE MÉDIA DE PICOS DE ACESSO À INTERNET NO PRIMEIRO TRIMESTRE DE 2013.	47
FIGURA 5 – CICLO DE VIDA DA INFORMAÇÃO DIGITAL.....	82
FIGURA 6 – O PAPEL CENTRAL DOS SOC NO PROCESSO DE RECUPERAÇÃO DA INFORMAÇÃO VISTO POR LANCASTER – IMAGEM REPRODUZIDA A PARTIR DA P. 2 DO SEU LIVRO <i>INDEXING AND ABSTRACTING IN THEORY AND PRACTICE</i> . .	94
FIGURA 7 – EXEMPLO DE NUVEM DE ASSUNTOS ELABORADA COM WORLD.NET A PARTIR DO TEXTO INTEGRAL DE <i>10 BEST APPLICATIONS DESIGN REPORT</i> POR JAKOB NIELSEN	99
FIGURA 8 – TAXONOMIA DE MODELOS DE RECUPERAÇÃO DA INFORMAÇÃO POR RICARDO BAEZA-YATES E BERTHIER RIBEIRO-NETO.....	107
FIGURA 9 - MODELO EM NINHO DAS ÁREAS DE BUSCA DA INFORMAÇÃO E DE PESQUISA DA INFORMAÇÃO DE WILSON .	110
FIGURA 10 – GRÁFICO RELATIVO À EVOLUÇÃO DOS REGISTOS DE REPOSITÓRIOS DE PORTUGAL NO OPENDOAR ENTRE 1 DE JANEIRO DE 2006 E 12 DE JANEIRO DE 2012	113
FIGURA 11 – MAPA DOS 45 REPOSITÓRIOS DIGITAIS EXISTENTES EM PORTUGAL CONTINENTAL E ARQUIPÉLAGOS DOS AÇORES E DA MADEIRA REGISTRADOS NO OPENDOAR PERTENCENTES A 42 ORGANIZAÇÕES (23/11/2013)....	115
FIGURA 12 - MAPA DOS 43 REPOSITÓRIOS DIGITAIS EXISTENTES EM PORTUGAL CONTINENTAL REGISTRADOS NO OPENDOAR (23/11/2013)	116
FIGURA 13 - MAPA DOS REPOSITÓRIOS DIGITAIS EXISTENTES NAS ZONAS NORTE E CENTRO DO TERRITÓRIO CONTINENTAL DE PORTUGAL REGISTRADOS NO OPENDOAR (23/11/2013)	117
FIGURA 14 - DISTRIBUIÇÃO GEOGRÁFICA DOS REPOSITÓRIOS DIGITAIS QUE SE LOCALIZAM NA ZONA DO GRANDE PORTO DO TERRITÓRIO CONTINENTAL DE PORTUGAL REGISTRADOS NO OPENDOAR (23/11/2013)	119
FIGURA 15 - MAPA DOS REPOSITÓRIOS DIGITAIS EXISTENTES NA ZONA SUL DO TERRITÓRIO CONTINENTAL DE PORTUGAL REGISTRADOS NO OPENDOAR (23/11/2013)	120
FIGURA 16 - DISTRIBUIÇÃO GEOGRÁFICA DOS REPOSITÓRIOS DIGITAIS QUE SE LOCALIZAM NA ZONA DA GRANDE LISBOA DO TERRITÓRIO CONTINENTAL DE PORTUGAL REGISTRADOS NO OPENDOAR (23/11/2013).....	121
FIGURA 17 – LOCALIZAÇÃO GEOGRÁFICA DO REPOSITÓRIUM DA UNIVERSIDADE DO MINHO NA ZONA NORTE DE PORTUGAL CONTINENTAL REGISTRADOS NO OPENDOAR (23/11/2013)	123
FIGURA 18 - REGISTO DO REPOSITÓRIUM DA UNIVERSIDADE DO MINHO NO OPENDOAR (23/11/2013)	124
FIGURA 19 - LOCALIZAÇÃO GEOGRÁFICA DOS REPOSITÓRIOS DIGITAIS DA UNIVERSIDADE DO PORTO NA ZONA NORTE DE PORTUGAL CONTINENTAL REGISTRADOS NO OPENDOAR (23/11/2013)	125
FIGURA 20 – REGISTO DO REPOSITÓRIO ABERTO DA UNIVERSIDADE DO PORTO NO OPENDOAR (23/11/2013).....	126

FIGURA 21 - LOCALIZAÇÃO GEOGRÁFICA DOS REPOSITÓRIOS DIGITAIS DA UNIVERSIDADE DE AVEIRO NA ZONA CENTRO DE PORTUGAL CONTINENTAL REGISTRADOS NO OPENDOAR (23/11/2013)	127
FIGURA 22 - REGISTO DO REPOSITÓRIO INSTITUCIONAL DA UNIVERSIDADE DE AVEIRO (RIA) NO OPENDOAR (23/11/2013)	128
FIGURA 23 – QUADRO SOBRE MÉTODOS DE INVESTIGAÇÃO POR FONTE DOS DADOS <i>VERSUS</i> TIPO DE ABORDAGEM <i>VERSUS</i> CONTEXTO DE USO DO PRODUTO REPRODUZIDO A PARTIR DO <i>SITE</i> DE JAKOB NIELSEN	134
FIGURA 24 – MODELO DO SISTEMA QUETTA	136
FIGURA 25 - MODELO DO SISTEMA QUETTA INICIANDO UM CONTADOR	137
FIGURA 26 – PÁGINA INICIAL DO <i>SITE</i> QUETTA	138
FIGURA 27 – PÁGINA DE LIGAÇÃO PARA O QUESTIONÁRIO INICIAL NO <i>SITE</i> QUETTA	140
FIGURA 28 - PÁGINA DO QUESTIONÁRIO INICIAL, OU QUESTIONÁRIO PRÉVIO AO TESTE DE USABILIDADE, NO <i>SITE</i> QUETTA (PARTE 1)	141
FIGURA 29 - PÁGINA DO QUESTIONÁRIO INICIAL, OU QUESTIONÁRIO PRÉVIO AO TESTE DE USABILIDADE, NO <i>SITE</i> QUETTA (PARTE 2)	141
FIGURA 30 - PÁGINA DO QUESTIONÁRIO INICIAL, OU QUESTIONÁRIO PRÉVIO AO TESTE DE USABILIDADE, NO <i>SITE</i> QUETTA (PARTE 3)	142
FIGURA 31 - PÁGINA DO QUESTIONÁRIO INICIAL, OU QUESTIONÁRIO PRÉVIO AO TESTE DE USABILIDADE, NO <i>SITE</i> QUETTA (PARTE 4)	142
FIGURA 32 - PÁGINA DO QUESTIONÁRIO INICIAL, OU QUESTIONÁRIO PRÉVIO AO TESTE DE USABILIDADE, NO <i>SITE</i> QUETTA (PARTE 5 E ÚLTIMA)	143
FIGURA 33 - INSTRUÇÕES PARA O TESTE 1 NA IU DO <i>EYETRACKER</i>	144
FIGURA 34 – BIBLIOTECA DIGITAL SINBAD DA UNIVERSIDADE DE AVEIRO (TESTE 1)	145
FIGURA 35 - PÁGINA DE LIGAÇÃO PARA O “QUESTIONÁRIO POSTERIOR 1 – SINBAD” NO <i>SITE</i> QUETTA.....	145
FIGURA 36 - PÁGINA DO QUESTIONÁRIO POSTERIOR 1 - SINBAD, OU QUESTIONÁRIO POSTERIOR AO TESTE DE USABILIDADE 1, NO <i>SITE</i> QUETTA (PARTE 1)	146
FIGURA 37 - PÁGINA DO QUESTIONÁRIO POSTERIOR 1 - SINBAD, OU QUESTIONÁRIO POSTERIOR AO TESTE DE USABILIDADE 1, NO <i>SITE</i> QUETTA (PARTE 2)	146
FIGURA 38 - PÁGINA DO QUESTIONÁRIO POSTERIOR 1 - SINBAD, OU QUESTIONÁRIO POSTERIOR AO TESTE DE USABILIDADE 1, NO <i>SITE</i> QUETTA (PARTE 3)	147
FIGURA 39 - PÁGINA DO QUESTIONÁRIO POSTERIOR 1 - SINBAD, OU QUESTIONÁRIO POSTERIOR AO TESTE DE USABILIDADE 1, NO <i>SITE</i> QUETTA (PARTE 4)	147
FIGURA 40 - INSTRUÇÕES PARA O TESTE 2 NA IU DO <i>EYETRACKER</i>	148
FIGURA 41 – REPOSITÓRIO ABERTO DA UNIVERSIDADE DO PORTO (TESTE 2)	149
FIGURA 42 - PÁGINA DE LIGAÇÃO PARA O “QUESTIONÁRIO POSTERIOR 2 – REPOSITÓRIO ABERTO U.PORTO” NO <i>SITE</i> QUETTA	149
FIGURA 43 - INSTRUÇÕES PARA O TESTE 3 NA IU DO <i>EYETRACKER</i>	150
FIGURA 44 - REPOSITÓRIO TEMÁTICO DA UNIVERSIDADE DO PORTO (TESTE 3).....	151
FIGURA 45 - PÁGINA DE LIGAÇÃO PARA O “QUESTIONÁRIO POSTERIOR 3 – REPOSITÓRIO TEMÁTICO U.PORTO” NO <i>SITE</i> QUETTA	151
FIGURA 46 - INSTRUÇÕES PARA O TESTE 4 NA IU DO <i>EYETRACKER</i>	152
FIGURA 47 – MEMÓRIAS DE ÁFRICA, DA FUNDAÇÃO PORTUGAL-ÁFRICA, MANTIDA PELA UNIVERSIDADE DE AVEIRO E PELO CENTRO DE ESTUDOS SOBRE ÁFRICA E DO DESENVOLVIMENTO (TESTE 4).....	153
FIGURA 48 - PÁGINA DE LIGAÇÃO PARA O “QUESTIONÁRIO POSTERIOR 4 – MEMÓRIAS DE ÁFRICA” NO <i>SITE</i> QUETTA	154
FIGURA 49 - INSTRUÇÕES PARA O TESTE 5 NA IU DO <i>EYETRACKER</i>	154
FIGURA 50 – REPOSITÓRIUM DA UNIVERSIDADE DO MINHO (TESTE 5)	155
FIGURA 51 - PÁGINA DE LIGAÇÃO PARA O “QUESTIONÁRIO POSTERIOR 5 – REPOSITORIUM” NO <i>SITE</i> QUETTA	156
FIGURA 52 – <i>EYETRACKER</i> DA MARCA TOBII (MODELO T120 EYE TRACKER) – PORMENOR DO CANTO SUPERIOR DIREITO – FOTO DE ANABELA SERRANO	158
FIGURA 53 - <i>EYETRACKER</i> DA MARCA TOBII (MODELO T120 EYE TRACKER) – VISTA LATERAL – FOTO DE ANABELA SERRANO.....	159
FIGURA 54 - <i>EYETRACKER</i> DA MARCA TOBII (MODELO T120 EYE TRACKER) – VISTA FRONTAL 1 – FOTO DE ANABELA SERRANO.....	160

FIGURA 55 - <i>EYETRACKER</i> DA MARCA TOBII (MODELO T120 EYE TRACKER) – VISTA FRONTAL 2 – FOTO DE ANABELA SERRANO	161
FIGURA 56 – DISPOSIÇÃO DOS ELEMENTOS DO EQUIPAMENTO E POSIÇÃO DA CADEIRA DO PARTICIPANTE FRENTE AO <i>EYETRACKER</i> DA MARCA TOBII (MODELO T120 EYE TRACKER) – FOTO DE ANABELA SERRANO	162
FIGURA 57 – EXEMPLO DE UM MAPA DE FIXAÇÕES OU <i>GAZE PLOT</i> COM AS FIXAÇÕES DE VÁRIOS PARTICIPANTES SOBREPOSTAS E DIFERENCIADAS PELA ATRIBUIÇÃO DE UMA COR DIFERENTE A CADA UTILIZADOR	167
FIGURA 58 – EXEMPLO DE <i>HEATMAP</i> POR CONTAGEM DE CLIQUES	168
FIGURA 59 - EXEMPLO DE <i>HEATMAP</i> POR DURAÇÃO ABSOLUTA	168
FIGURA 60 – GRÁFICO DA ATIVIDADE DOS PARTICIPANTES NA UNIVERSIDADE DE AVEIRO	173
FIGURA 61 - GRÁFICO DO GRUPO ETÁRIO DOS PARTICIPANTES	174
FIGURA 62 - GRÁFICO DO NÍVEL DE HABILITAÇÕES ACADÉMICAS DOS PARTICIPANTES	175
FIGURA 63 - GRÁFICO RELATIVO À PESQUISA POR ASSUNTO EM PLATAFORMAS DIGITAIS PELOS PARTICIPANTES	176
FIGURA 64 - GRÁFICO RELATIVO À FREQUÊNCIA DE ACESSO ÀS PLATAFORMAS DIGITAIS PELOS PARTICIPANTES	177
FIGURA 65 - GRÁFICO RELATIVO À FREQUÊNCIA DE PESQUISA POR ASSUNTO COMPARATIVAMENTE À PESQUISA POR AUTOR E TÍTULO PELOS PARTICIPANTES	177
FIGURA 66 – GRÁFICO COMPARATIVO ENTRE A FREQUÊNCIA DA PESQUISA POR ASSUNTO E A ATIVIDADE NA UNIVERSIDADE DE AVEIRO	178
FIGURA 67 - GRÁFICO COMPARATIVO ENTRE OS DIFERENTES REPOSITÓRIOS DIGITAIS QUANTO À FACILIDADE DE ACESSO À INFORMAÇÃO	179
FIGURA 68 - GRÁFICO COMPARATIVO ENTRE OS DIFERENTES REPOSITÓRIOS DIGITAIS QUANTO À EXISTÊNCIA/LOCALIZAÇÃO DE AUXÍLIO À PESQUISA POR ASSUNTO NA IU	180
FIGURA 69 - GRÁFICO COMPARATIVO ENTRE OS DIFERENTES REPOSITÓRIOS DIGITAIS QUANTO À IDENTIFICAÇÃO DE SOC NA IU	181
FIGURA 70 - GRÁFICO COMPARATIVO ENTRE OS DIFERENTES REPOSITÓRIOS DIGITAIS QUANTO AO GRAU DE SATISFAÇÃO DO UTILIZADOR	182
FIGURA 71 - GRÁFICO COMPARATIVO ENTRE OS DIFERENTES REPOSITÓRIOS DIGITAIS QUANTO AO PAPEL FACILITADOR DA PRESENÇA DOS SOC NA PESQUISA POR ASSUNTO	184
FIGURA 72 - MAPA DE FIXAÇÕES DOS PARTICIPANTES COM 40-49 ANOS (TESTE 1)	185
FIGURA 73 - MAPA DE FIXAÇÕES DOS PARTICIPANTES COM DOUTORAMENTO, 40-49 ANOS, E DOCENTES OU INVESTIGADORES (TESTE 1)	186
FIGURA 74 – MAPA DE FIXAÇÕES DOS PARTICIPANTES COM HABILITAÇÕES ACADÉMICAS AO NÍVEL DO ENSINO SECUNDÁRIO (TESTE 1)	186
FIGURA 75 - MAPA DE FIXAÇÕES DOS PARTICIPANTES COM O GRAU DE MESTRE (TESTE 1)	187
FIGURA 76 - MAPA DE FIXAÇÕES DOS PARTICIPANTES COM O GRAU DE LICENCIADO (TESTE 1)	187
FIGURA 77 - MAPA DE FIXAÇÕES DOS PARTICIPANTES COM O GRAU DE LICENCIADO, EXCLUINDO O ID 19 (TESTE 1)	188
FIGURA 78 - MAPA DE FIXAÇÕES DOS PARTICIPANTES COM 18-29 ANOS (TESTE 1)	188
FIGURA 79 - MAPA DE FIXAÇÕES DOS PARTICIPANTES COM 18-29 ANOS, EXCLUINDO O PARTICIPANTE COM O ID 19 (TESTE 1)	189
FIGURA 80 - MAPA DE FIXAÇÕES DOS PARTICIPANTES QUE NÃO IDENTIFICARAM NENHUM SOC (TESTE 1)	189
FIGURA 81 - REPOSITÓRIO ABERTO DA UNIVERSIDADE DO PORTO (TESTE 2)	190
FIGURA 82 - REPOSITÓRIO TEMÁTICO DA UNIVERSIDADE DO PORTO (TESTE 3)	190
FIGURA 83 - MEMÓRIAS DE ÁFRICA MANTIDA PELA UNIVERSIDADE DE AVEIRO (TESTE 4) – EXEMPLO 1	191
FIGURA 84 - MEMÓRIAS DE ÁFRICA MANTIDA PELA UNIVERSIDADE DE AVEIRO (TESTE 4) – EXEMPLO 2	191
FIGURA 85 - MEMÓRIAS DE ÁFRICA MANTIDA PELA UNIVERSIDADE DE AVEIRO (TESTE 4) – EXEMPLO 3	191
FIGURA 86 - REPOSITÓRIUM DA UNIVERSIDADE DO MINHO (TESTE 5)	192
FIGURA 87 – GRÁFICO DOS TEMPOS, EM MINUTOS, DE REALIZAÇÃO DOS TESTES DE USABILIDADE COM <i>EYETRACKING</i> 1 A 5, EXCLUINDO OS TEMPOS DE CALIBRAGEM 1 A 5 E DE INSTRUÇÕES PARA O TESTE 1 A 5, MAIS O TEMPO TOTAL DE REALIZAÇÃO DO ESTUDO COM O SISTEMA QUETTA POR PARTICIPANTE	193
FIGURA 88 – GRÁFICO DOS TEMPOS, EM MINUTOS, DE REALIZAÇÃO DOS TESTES DE USABILIDADE COM <i>EYETRACKING</i> 1 A 5, EXCLUINDO OS TEMPOS DE CALIBRAGEM 1 A 5 E DE INSTRUÇÕES PARA O TESTE 1 A 5, POR PARTICIPANTE.	193

Glossário de acrónimos e siglas

DLF – Digital Library Federation

ERP – Enterprise Resource Planning

EUA – Estados Unidos da América

I&D – Investigação e Desenvolvimento

IS&R – *Information Seeking & Retrieval*

IU – Interface do Utilizador

mbps – megabit por segundo

OPAC – Online Public Access Catalog

OpenDOAR – Directory of Open Access Repositories

PI – Pesquisa da Informação

QUETTA – Questionários, Usabilidade com EyeTracking e Think-Aloud

RCAAP – Repositório Científico de Acesso Aberto de Portugal

RI – Recuperação da Informação

ROAR - Registry of Open Access Repositories

SAP – Systems, Applications and Products

SI – Sistemas de Informação

SOC – Sistemas de Organização do Conhecimento

SRI – Sistemas de Recuperação da Informação

TI – Tecnologias da Informação

«Where is the life we have lost in living?

Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in information?»

T.S. Eliot (1934), “The Rock”

Introdução

1. Contextualização do projeto de investigação

Atualmente, a sociedade humana estrutura-se em torno da informação e do conhecimento. Ambos têm uma relação direta e intrínseca com o Homem, perpassando toda a existência humana e a história da humanidade. Não obstante, as sociedades de outros séculos estruturaram-se em torno de outros fatores de desenvolvimento ou de organização mais prementes para sociedades menos evoluídas em que os níveis de conforto socioeconómico eram bem menores. A informação decorre da ação humana e constitui suporte para a mesma. Podemos, inclusive, no estágio de desenvolvimento em que se encontra a neurociência, afirmar que o cérebro humano¹ constitui, em si-mesmo, um suporte de informação.²

¹ O cérebro humano é cerca de trinta vezes mais potente do que o mais rápido supercomputador da actualidade - o Tianhe-2 - cuja velocidade de processamento é de 33.86 petaflop/s. Apresentado na primeira semana de Junho de 2013 pela China, e em testes na National University for Defense Technology até ser montado no National Supercomputer Center em Guangzhou, o Tianhe-2 era o supercomputador mais avançado do mundo em 2 de Setembro de 2013. Para simular a velocidade de processamento da atividade cerebral humana será necessário cerca de 1 exaflop/s. Dados fornecidos pelo IEEE na sua página no Facebook:
<https://www.facebook.com/photo.php?fbid=551395268248446&set=a.176108879110422.62121.176104589110851&type=1&theater>

A Intel estima conseguir ter um computador com a velocidade de processamento de 1 exaflop/s por volta de 2018: <http://www.kurzweilai.net/intel-plans-exaflops-supercomputer-by-2018>, o que talvez venha a permitir cumprir a visão de alguns cientistas da inteligência artificial e da neurociência computacional, defendida pelo famoso físico inglês Stephen Hawking, de copiar o cérebro humano para dentro de um computador permitindo, mesmo, a vida para além da morte.
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4531462>

A análise, tratamento e organização da informação têm sido levadas a cabo pelo Homem desde há muitos séculos, conhecendo-se vestígios da existência de sistemas de informação (SI) analógicos nas civilizações pré-clássicas dos vales do Tigre e do Eufrates e do vale do Nilo. As trocas comerciais e as viagens favoreciam a troca de informação e tornavam imperiosa a gestão da informação recebida e, sobretudo, daquela que era criada para apoio às atividades económicas. Porém, a difusão da informação fazia-se de forma lenta e irregular, quando não era ocasional, devido às características rudimentares dos meios de reprodução e de comunicação. Antes do aparecimento da escrita e dos primeiros alfabetos, a informação era difundida pela comunicação não-verbal (pictórica ou graficamente³) e pela comunicação verbal oral (transmissão oral da história, da cultura e do imaginário das tribos, sociedades, civilizações). A comunicação verbal escrita terá surgido cerca de 3000 a.C. e foi manuscrita, assente em vários suportes (tabuinhas de argila e de cera, papiro, pergaminho e papel), até ao aparecimento da imprensa de Gutenberg no século XV. A difusão da informação fez-se de forma lenta, e algo entrecortada, ao longo de alguns milénios. Paralelamente, a comunicação do conhecimento processava-se de forma ainda mais lenta, dado que os locais de ensino/aprendizagem eram poucos e reservados à elite (nobreza e clero) ou a parte dela, em geral, sob a forma de lições particulares. A escola pública surge, apenas, no século VIII da era de Cristo no império carolíngio (França e Alemanha), por iniciativa legislativa de Carlos Magno. Os níveis de literacia de então eram muitíssimo baixos, mesmo no seio do clero, e Carlos Magno legislou sobre educação, entre muitas outras áreas, criando uma escola para todos, gratuita, à qual tiveram acesso até mesmo os filhos dos servos. Este terá sido o impulso necessário para que surgissem, na Europa do século XII, as primeiras universidades. A transmissão do conhecimento nas universidades apoiar-se-ia, nos

² John von Neumann chegou à conclusão de que o cérebro armazena centenas de MB de informação por segundo no seu estudo “Die Rechenmaschine und das Gehirn” (O Computador e o Cérebro) em 1991. William Bialek, et al., publicou na revista Science, no mesmo ano, “Reading a Neural Code” onde estuda o fluxo de transmissão da informação veiculada pelo nervo ótico no cérebro a taxa da informação transmitida pelos neurónios. O livro “Spikes” da autoria de Fred Rieke, e mais alguns artigos recentemente publicados elucidam melhor esta questão.

³ As gravuras rupestres, por exemplo, são vestígios de uma forma de expressão que pode ser vista como arte, mas, talvez, também como uma forma de comunicação.

primeiros séculos, em códices que replicavam obras clássicas, frequentemente, com falhas e imprecisões de tradução.

Após 1456, com a primeira edição de uma obra – a Bíblia – por impressão em papel com caracteres móveis metálicos, na Europa na tipografia de Gutenberg, a comunicação do conhecimento acelera-se subitamente. O tempo e os custos de reprodução das obras baixam vertiginosamente, passando a utilizar-se o papel (puramente vegetal) em vez do pergaminho como suporte preferencial da informação. O códice é “substituído” pelo livro (incunábulo até 1500). Os séculos XV e XVI constituem, igualmente, os séculos dos Descobrimentos e do avanço tecnológico e científico. O conhecimento científico conhece um desenvolvimento sem precedentes, bem como as obras que o difundem. Nas universidades começam a ser ensinadas novas disciplinas e surgem novos cursos voltados para as ciências exatas e aplicadas. O número de letrados aumenta e as bibliotecas das universidades crescem competindo com as particulares e suplantando-as em atualidade, cariz científico e acessibilidade à informação. O século XVIII vê desenvolverem-se as ideias iluministas trazidas pelos estrangeirados, e surgem a primeira grande obra de referência – “L’Encyclopédie” de Diderot e D’Alembert⁴ - e as primeiras publicações periódicas com o aparecimento das gazetas – os primeiros jornais. São os primórdios da comunicação social. A Revolução Francesa prepara-se sob este fervilhar de ideias e difusão de informação que permitem, conjugados com o descontentamento do Terceiro Estado cada vez mais poderoso económica e culturalmente (a esmagadora maioria da população onde se incluíam os burgueses e a maioria dos letrados), a comunicação das ideias republicanas a toda a França e até ao resto da Europa e à América. No Novo Mundo, não tardaria a surgir outra nação independente e republicana – os Estados Unidos da América (EUA) -, que muito iria impulsionar o avanço tecnológico e científico nos séculos seguintes. Em Inglaterra, ainda no século XVIII, o avanço tecnológico conduziu à “Revolução Industrial” e a um novo paradigma de desenvolvimento – toda uma nova estrutura que substituirá uma

⁴ Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers, par une société de gens de lettres / mis en ordre et publié par M. Diderot et quant à la partie mathématique par M. D’Alembert. À Paris : Chez Briasson : David : Le Breton : Durand, 1751-1780. Acessível na Internet: «<http://porbase.bn.pt/>».

sociedade marcadamente rural – que iria dominar a sociedade ocidental durante mais de século e meio.

O século XIX (e a primeira metade do século XX) pode ser visto como o século do desenvolvimento tecnológico e científico voltado para a produção industrial. É nesta altura que surgem as bibliotecas públicas com um grande impulso dado quer em França quer nos EUA, quer em Portugal em consequência da perseguição do Marquês de Pombal aos jesuítas e das revoltas liberais. A Real Biblioteca Pública da Corte é criada, em Lisboa, ainda em finais do século XVIII (em 29 de Fevereiro de 1796), com o objetivo de permitir o acesso aos documentos ao público em geral. No Porto, no início do século XIX, Alexandre Herculano é o bibliotecário responsável por documentação proveniente de bibliotecas particulares eclesiásticas sediadas em conventos e mosteiros espoliados e/ou abandonados (incluindo Santa Cruz em Coimbra) e inicia a sua catalogação para a disponibilizar ao público – a 9 de Julho de 1833 era criada a Real Biblioteca Pública da Cidade do Porto que se tornaria no segundo maior fundo documental do país atingindo perto de 100 000 exemplares⁵ logo nos primeiros anos, e cerca de 1 300 000 documentos impressos na atualidade⁶. Nos EUA, a Library of Congress, criada em 1800 e aberta ao público 97 anos depois (em 1 de Novembro de 1897), cresce aceleradamente e, no início do século XX, conta já com cerca de 1 000 000 de volumes⁷. Os meios de comunicação social sofrem o impacto do desenvolvimento tecnológico do século XIX, bem patente na

⁵ O fundo documental inicial seria composto por cerca de 36000 volumes impressos e 300 códices. Em 1842, à data da abertura da biblioteca, contaria já com cerca de 97000 volumes. A biblioteca começou por estar instalada no Hospício dos Franciscanos na Praça da Cordoaria em 1841, tendo mudado, em 1842, para o edifício do Convento de Santo António da Cidade em frente ao Jardim de S. Lázaro onde, ainda, hoje se encontra sediada. A abertura oficial data de 4 de Abril de 1842. Dados retirados da ficha de descrição do monumento Convento de Santo António da Cidade constante no website do Instituto da Habitação e da Reabilitação Urbana (IHRU). Acessível na Internet:

«<http://www.monumentos.pt/Monumentos/usercontrols/Pdf.aspx?Ipas=5418&&Type=FICHAIPA>». Foram, ainda, retirados dados de uma página da wikipédia relativa à Biblioteca Pública Municipal do Porto (BPMP). Acessível na Internet:

«http://pt.wikipedia.org/wiki/Biblioteca_P%C3%BAblica_Municipal_do_Porto».

⁶ Dados referidos por Júlio Manuel Rodrigues Costa, Chefe de Divisão Municipal de Coleções e Desenvolvimento da BPMP, em 27 de Outubro de 2005, numa comunicação apresentada nas comemorações do bicentenário da Biblioteca Pública de Évora. Acessível na Internet:

«http://www.evora.net/BPE/2005Bicentenario/dias/27_out05/textos/julio.pdf».

⁷ No início do século XXI, são já 144 000 000 de exemplares, de acordo com dados disponibilizados no website da Library of Congress. Acessível na Internet: «<http://www.loc.gov/about/history.html>».

“Exposition Universelle” de 1900 em Paris, e os jornais passam a ter fotografias a ilustrar os textos das colunas, em vez de gravuras, e surge a rádio com as primeiras emissões contínuas (1916 nos EUA) e, mais tarde, a televisão (1928 nos EUA). Ambos os meios de comunicação tiveram uma fase experimental (fase que, ainda, teve início no século XIX no caso da rádio), não havendo consenso no que diz respeito às origens da rádio, e difundiram-se rapidamente por todo o globo. As emissões contínuas de emissoras de rádio cimentaram-se nos anos 20 e 30 do século XX, e a televisão comercial teve a sua maior expansão nos anos 40 a 60, embora ainda hoje haja nações a inaugurarem a sua primeira estação televisiva⁸.

No século XX surge uma sociedade estruturada em torno da indústria, sobretudo na primeira metade, e em torno da informação nas últimas duas décadas, mas bem poderia ser considerado o século da comunicação social. A informação cresce, exponencialmente, e, ao longo de todo o século XX, o conhecimento científico é fomentado e comunicado a uma escala inédita. À data do final da 2ª Grande Guerra Mundial, Julho de 1945⁹, o desenvolvimento tecnológico era, segundo Vannevar Bush, suficiente para melhorar em larga escala as condições de vida da população mundial. O facto de a Informática ter surgido logo após o final da guerra só confirma esta afirmação. Professor do MIT, com as mais altas responsabilidades a nível científico¹⁰ nos EUA durante a 2ª Grande Guerra Mundial, Bush influenciou a evolução tecnológica, científica e social dos últimos 65 anos ao publicar, sob o título “As we may think” na *The Atlantic Monthly* a 1 de Julho de 1945, uma visão para o futuro que o próprio tinha entregado sob a forma de relatório ao Presidente Roosevelt. Nesse artigo que foi publicado numa versão de divulgação ao grande público (a 10 de Setembro do mesmo ano) na revista *Life*, podem ver-se os traços,

⁸ O principado do Liechtenstein em 2008.

⁹ O artigo de Vannevar Bush foi publicado no dia 1 de Julho de 1945, e é essa data que aqui é invocada no seu contexto histórico. O final da 2ª Grande Guerra Mundial decorreu de Maio a Setembro de 1945, começando com a rendição alemã às Forças Aliadas em 7 de Maio de 1945 (Charles De Gaulle anunciou a rendição alemã no dia seguinte, 8 de Maio de 1945, que ficou conhecido como o “Dia da Vitória na Europa”) e terminando com a rendição do Japão assinada em 2 de Setembro de 1945 que ficou conhecido como o “Dia da Vitória sobre o Japão”.

http://www.nationalgeographic.com/pearlharbor/history/wwii_timeline.html

<http://www.historyplace.com/worldwar2/timeline/ww2time.htm#1945>

¹⁰ Diretor do “Office of Scientific Research and Development”.

não só do hipertexto e da Internet¹¹ como é mais usual focar, mas, também, dos arquivos e das bibliotecas digitais na descrição que o autor faz do dispositivo “Memex”. A maior preocupação de Bush era a impossibilidade de, então, recuperar toda a informação existente para uma dada necessidade de informação e a alta probabilidade de equipas de investigação dedicarem tempo e recursos a projetos que poderiam estar a ser desenvolvidos, em simultâneo, em locais diferentes do planeta sem que cada equipa conhecesse o trabalho da outra. Perante uma já perceptível “explosão da informação”, esta realidade resultava em enorme desperdício de financiamentos que poderia ser evitado se fosse possível ter numa secretária “eletrónica” (“Memex”) o nosso arquivo e a nossa biblioteca e ligá-los aos arquivos e bibliotecas de outros cientistas por meio de uma rede de dados. Esta visão tem vindo a tornar-se realidade ao longo das últimas décadas e teve implicações no aparecimento da Ciência da Informação e da recuperação da informação como área de investigação. A Internet, enquanto rede de comunicação de dados, constitui um suporte de informação e um meio de comunicação e de publicação de dimensões faraónicas. Recuperar informação com eficácia continua a ser um desafio para os cientistas do século XXI. Mas, é-o, igualmente, para todo o resto da população mundial. Todos somos produtores e consumidores de informação – encontramos-nos numa sociedade global designada de “sociedade da informação”. Usamos informação com uma enorme frequência: a que produzimos, a que recebemos e aquela que pesquisamos e conseguimos recuperar. Pesquisar informação tornou-se numa rotina sem horários e sem deslocações graças aos SI digitais disponíveis na *World Wide Web*. Porém, o volume de informação disponível é avassalador, pelo que a satisfação das nossas necessidades de informação depende, em muito, da forma como a interação homem-máquina decorre e não, apenas, da eficácia dos sistemas de recuperação da informação. A acessibilidade da informação depende do *design* da interface do utilizador (IU) de um dado sistema de informação e do seu conteúdo. O excesso de informação e a necessidade constante de a recuperar fazem com que a pesquisa por assunto ganhe cada vez maior importância. Há, no entanto, que verificar se a pesquisa por assunto pode ser sempre realizada nas melhores condições, isto é,

¹¹ Tim Berners-Lee, que criou a *World Wide Web* em 1989, admite ter sido inspirado pela leitura desse artigo de Bush na *Life* quando ainda era adolescente.

se os SI digitais disponibilizam o acesso a sistemas de organização de conhecimento e de que forma o fazem. Para isso será necessário avaliar os SI digitais que disponibilizam informação técnico-científica como é o caso das bibliotecas digitais e dos arquivos digitais.

A escolha do tema desta tese recai, pelos motivos acima enunciados, sobre a qualidade dos SI digitais referidos, habitualmente, pelas designações de “arquivos digitais”, de “repositórios institucionais” e de “bibliotecas digitais”.

De facto, os arquivos e as bibliotecas digitais constituem uma realidade recente que embora esteja, ainda, em fase de crescimento, se encontra, igualmente, num momento de reequacionamento em termos científicos, pelo que é premente proceder a uma avaliação daqueles que já se encontram acessíveis na Internet. Acresce referir que, a nível internacional, são ainda poucos os estudos desenvolvidos nesse sentido não existindo, ainda, normas publicadas, suficientemente abrangentes, pelo que poderá ser necessário definir um modelo de avaliação aplicável a estes tipos de SI, de forma generalizada.

2. Objeto de estudo e objetivos de investigação

O objeto de estudo desta tese são os SI do tipo “biblioteca digital” e “arquivo digital” disponíveis na Internet com origem em organizações sediadas em Portugal.¹²

O projeto de investigação desenvolvido teve por objetivo efetuar uma avaliação da qualidade dos SI do tipo arquivo digital e biblioteca digital, em termos de acessibilidade da informação, fazendo o ponto da situação do seu desenvolvimento, atualmente, em Portugal.

¹² É necessário definir o conceito e as diferentes tipologias de sistema de informação possíveis. Definir biblioteca digital e arquivo digital bem como repositório digital. Elucidar outros aspetos acerca de recursos de informação disponíveis na Internet.

A acessibilidade da informação pode ser determinada avaliando as formas de recuperação da informação, nas quais se incluem os designados sistemas de organização do conhecimento (SOC) cuja acessibilidade pode ser analisada por intermédio da IU, nestes SI digitais. Será necessário equacionar a necessidade de criar um modelo de avaliação de arquivos e bibliotecas digitais no contexto de avaliação atrás indicado – a acessibilidade da informação e a recuperação da informação. Para atingir o objetivo principal, será necessário estudar os comportamentos informacionais dos utilizadores deste tipo de SI no que diz respeito à sua interação com a IU, bem como a eficácia dos sistemas de recuperação da informação (SRI) em função da presença ou inexistência de SOC.

3. Questão de investigação e hipóteses

As quatro questões que desencadearam este projeto de investigação, e cujas respostas se procuram alcançar no decurso do mesmo, são as seguintes:

Questão 1 - Os SOC quando presentes nos SI digitais do tipo arquivo ou biblioteca estão acessíveis?

Questão 2 - A utilização de SOC nos SI digitais facilita a recuperação da informação por assunto?

Questão 3 - O conhecimento prévio dos SOC por parte do utilizador influencia positivamente o sucesso na pesquisa por assunto, diminuindo o tempo de obtenção de resultados satisfatórios?

Questão 4 - A presença dos SOC nos SI digitais influencia o grau de satisfação do utilizador com o SI digital?

Cada uma destas questões conduz à formulação de hipóteses que necessitarão de ser confirmadas, ou não, no final da investigação, tais como:

Hipótese 1 – Os SOC encontram-se acessíveis na página inicial do SI digital;

Hipótese 2 – A utilização de SOC facilita a recuperação da informação por assunto no SI digital;

Hipótese 3 – O conhecimento prévio dos SOC por parte do utilizador influencia positivamente o sucesso na pesquisa por assunto, diminuindo o tempo de obtenção de resultados satisfatórios;

Hipótese 4 – A presença dos SOC influencia positivamente o grau de satisfação do utilizador com o SI digital.

Mas, para poder atingir os objetivos da investigação respondendo às questões de investigação, e confirmando ou não as hipóteses acima formuladas, há que definir qual ou quais o(s) método(s) de investigação a adotar.

4. Método de investigação

Para o desenvolvimento do projeto de investigação aqui descrito foi, inicialmente, após uma revisão da literatura científica publicada, selecionada uma combinação de abordagens quantitativas e qualitativas adequadas à avaliação de bibliotecas digitais: os testes de usabilidade, o inquérito, a análise dos percursos seguidos pelos utilizadores em cada sessão - *web log analysis* – e o *think-aloud*. Estas abordagens contextualizam-se nos seguintes métodos de investigação: experimental, inquérito, análise de conteúdo e observação.

Os testes de usabilidade permitem avaliar o acesso à informação e a interface do utilizador sob a perspetiva do utilizador. Pelo facto de a localização dos conteúdos numa página web poder ser determinante para a acessibilidade dos mesmos, optou-se pelos testes de usabilidade com *eyetracking* como técnica de investigação ideal, não obstante o carácter dispendioso do equipamento necessário para o efeito, o que poderia ser um obstáculo à realização do trabalho laboratorial. Não havendo disponibilidade de um laboratório equipado com *eyetracker* que permitisse o

cumprimento da calendarização prevista (2010) no plano do projeto de investigação (o plano de doutoramento) chegou a ponderar-se a substituição deste método por estudos de usabilidade *online* de larga escala. Este tipo de estudo de usabilidade permitiria abranger uma larga amostra populacional, mas traria o inconveniente de não permitir a observação direta de cada um dos participantes, ainda que pudesse ser possível cruzar os dados do estudo com os do inquérito e, eventualmente, com os da *web log analysis*.

A observação em ambiente controlado dos participantes permite o cruzamento dos dados recolhidos através dos vários métodos de investigação a utilizar com maior rigor. Permite, igualmente, a criação de um modelo integrado de estudo do comportamento do utilizador, o que confere maior rigor aos testes a realizar direcionando-os para o estudo do ponto de vista do utilizador e garantindo o controlo de todas as fases do estudo.

Assim, à medida que o projeto avançou tornou-se evidente a necessidade de integrar as várias abordagens num novo sistema criado, propositadamente, para o projeto de investigação em causa.

Os testes de usabilidade foram realizados com *eyetracking*, em Maio de 2011, num *eyetracker* cedido pela Portugal Telecom SGPS SA (PT)¹³ ao Centro de Estudos das Tecnologias e Ciências da Comunicação (CETAC.MEDIA)¹⁴ que, por sua vez, o cedeu a este projeto de investigação por solicitação do Instituto de Engenharia Eletrónica e Telemática de Aveiro (IEETA)¹⁵. Estas duas unidades de investigação (o CETAC.MEDIA das Universidades do Porto e de Aveiro e o IEETA da Universidade de Aveiro) e aquela empresa (PT) forneceram os meios necessários para o avanço do projeto e para a realização, atempada (ainda que com mais de um semestre de atraso relativamente ao previsto), do trabalho laboratorial previsto no plano de doutoramento. O trabalho foi realizado nas instalações do CETAC.MEDIA no Departamento de Comunicação e Arte (DeCA) da Universidade de Aveiro.

¹³ Website da PT: <http://www.telecom.pt/InternetResource/PTSite/PT/>

¹⁴ Website do CETAC.MEDIA: <http://www.cetacmedia.org/>

¹⁵ Website do IEETA: http://wiki.ieeta.pt/wiki/index.php/Main_Page

5. Fontes de informação

Para o desenvolvimento deste projeto de investigação foram consultadas inúmeras fontes de informação de várias tipologias e em vários suportes quer em SI digitais quer em SI analógicos. Ressalve-se, no entanto, que a maioria das pesquisas foi conduzida em SI digitais. Nos motores de pesquisa, o Google foi o recurso preferencial tendo sido efetuado um rastreio diário para várias palavras-chave usando o serviço de alertas. Foram acedidos, igualmente, arquivos digitais, bibliotecas digitais, portais temáticos, bases de dados bibliográficas e em texto integral, Online Public Access Catalogs (OPAC), redes sociais científicas como a Mendeley e a Research Gate, sítios e páginas web de investigadores, blogs de investigadores, entre outros. Foram recuperados e gravados 1331 documentos, e foram acedidos e registados em perfil *online* cerca de 205 documentos no Google Books e cerca de 1100 na ACM. Segue-se uma relação não exaustiva das fontes de informação consultadas: Google Académico; Google Books; Scribd; ACM Digital Library; IEEE Computer Society Digital Library; Dissertation Abstracts; Proquest Dissertation & Theses – Proquest Digital Dissertations; UMI Dissertation Express; LISA; LISTA; LLIS; Librarians Digital Library; DELOS Network of Excellence on Digital Libraries; Scopus; Academic Search Complete; ERIC; EBSCO host; ISI Current Contents; ISI Web of Knowledge; ISI Web of Science; Compendex; INSPEC; Wilson SciTech; Wilson SocSci; catálogos de várias bibliotecas; portais e repositórios Open Access (DOAJ, Scielo, etc.); bases de dados em texto integral (ScienceDirect, Emerald, Springer Link, etc.); periódicos eletrónicos (Ariadne, DLib, Biblios, etc.).

6. Organização da tese

Esta tese encontra-se organizada em quatro capítulos sendo os dois primeiros relativos ao enquadramento teórico (capítulos 1 e 2) e os outros dois relativos ao

projeto de investigação aplicada realizado (capítulos 3 e 4). No capítulo 1 proceder-se-á a uma revisão da literatura sobre o tema da tese sob o título “A avaliação de arquivos e bibliotecas digitais: o estado da arte”, no capítulo 2 será realizada uma análise e definição concetual dos termos que são utilizados, nas fontes de informação analisadas, para enunciar os SI digitais que são objeto de estudo do projeto de investigação relatado nesta tese, bem como os aspetos teóricos relativos ao contexto de avaliação dos mesmos, sob o título “Os sistemas de informação digitais e o acesso à informação: conceitos e contextualização científica”. No capítulo 3, sob o título “O sistema QUETTA ou um Modelo Integrado de Estudo do Comportamento do Utilizador”, é descrito o modelo integrado de estudo do comportamento do utilizador criado no âmbito deste projeto de investigação e a forma como foi realizada a recolha de dados a partir daquele, após uma análise da forma como foi constituída a amostra dos arquivos e bibliotecas digitais a avaliar. Por último, no capítulo 4, são descritos os procedimentos de análise, tratamento e interpretação dos dados e são apresentados e discutidos os resultados da investigação, sob o título “Análise dos dados e discussão dos resultados da investigação”.

Capítulo 1

A avaliação de arquivos e bibliotecas digitais: o estado da arte

1.1. A avaliação de arquivos e bibliotecas digitais – estado da arte

Conforme foi referido na introdução, apesar do enorme crescimento do número de arquivos e bibliotecas digitais presentes na *World Wide Web*, os estudos referentes à avaliação da qualidade destes SI têm-se centrado nas bibliotecas digitais e são, ainda, em número reduzido – na sua maioria constituem estudos teóricos, contabilizando-se, apenas, cerca de centena e meia de estudos de investigação aplicada em todo o mundo no início do século XXI. A qualidade dos serviços prestados por estes SI permanece insondável na maioria dos casos, o que não permite aferir o valor e o impacto dos mesmos na sociedade atual nem criar modelos de desenvolvimento deste tipo de SI orientados para os utilizadores.

Segundo **Chowdhury (2003)**, as bibliotecas digitais são planeadas e desenvolvidas de modo a facilitarem o acesso à informação e a disponibilização de serviços aos utilizadores. Porém, nem sempre estes e outros objetivos intrínsecos ao seu desenvolvimento são devidamente atingidos, o que evidencia a necessidade de sujeitar estes projetos a estudos de avaliação, de forma a melhorar o seu desempenho e a permitir que surjam bibliotecas digitais cada vez mais eficientes que satisfaçam as necessidades de informação de um público-alvo cada vez mais difuso e heterogéneo. Em virtude de inúmeras características que não são partilhadas pelos serviços de informação analógicos, a avaliação deste tipo de SI digital é, eminentemente, complexa. A abrangência espacial e temporal que permite o acesso à informação 24 horas por dia, 365 dias no ano, a partir de qualquer ponto do planeta Terra onde se possa ter acesso à Internet, cria um público-alvo difícil de caracterizar e de satisfazer. Avaliar uma biblioteca digital implica avaliar processos caracterizados por interatividade entre o utilizador e a máquina/sistema, num contexto variado.

Marchionini (2000) refere que avaliar uma biblioteca digital é um pouco como julgar o quão bem-sucedido é um casamento, uma vez que o mesmo depende em parte do êxito individual dos parceiros e, por outro lado, das condições criadas pela união entre aqueles.

Saracevic (2000) afirma que tudo o que diz respeito a bibliotecas digitais é explosivo, tudo... exceto a avaliação. Segundo este autor, a literatura científica acerca da avaliação de bibliotecas digitais é, basicamente, de dois tipos: a metaliteratura que não contém dados e a literatura objeto que contém dados, sendo esta última muito menos significativa em termos de produção científica. A avaliação de bibliotecas digitais tem-se estruturado em torno de vários aspetos: o que é avaliado - entidades ou processos; o contexto de avaliação; os critérios de avaliação; a metodologia de avaliação; e os resultados da investigação. Sendo o perfil do utilizador, deste tipo de SI de acesso remoto, difícil de identificar, é urgente a realização de estudos de avaliação destes SI digitais que incidam sobre dados (daí que este seja um projeto de investigação aplicada) e sobre a perspetiva do utilizador (este projeto incide sobre a acessibilidade da informação).

De acordo com **Xie (2006)**, a identificação de critérios de avaliação pouco tem evoluído, sobretudo, no que diz respeito à perspetiva do utilizador. Os modelos de avaliação são criados sob a influência dos pressupostos ou arquétipos mentais dos seus autores. Por outro lado, a história da avaliação de bibliotecas (analógicas, híbridas ou digitais) evidencia uma prática inicial da avaliação de bibliotecas digitais muito influenciada por perspetivas de gestão de serviços de informação analógicos ou híbridos. É, sobretudo, no campo da investigação que os autores se conseguem destacar um pouco mais de arquétipos mentais influenciados por estruturas sociais e profissionais decorrentes de práticas quotidianas associadas a um modelo de desenvolvimento social, económico, político e cultural analógico. Torna-se, portanto, evidente que a avaliação de bibliotecas digitais tem raízes na avaliação de bibliotecas analógicas que passou por várias fases desde o início dos anos 30 do século XX nos EUA até ao início do século XXI. A necessidade de avaliar serviços de informação ganhou expressão nos últimos quarenta anos mercê da publicação de vários estudos acerca de avaliação de desempenho, de medidas e indicadores, de avaliação da qualidade, etc., em bibliotecas, serviços e sistemas de informação. A terminologia de então não era uniforme, e continuava a não o ser, ainda, há alguns anos atrás aquando da publicação das principais normas internacionais. Para além de não ser uniforme, cada termo é utilizado para representar conceitos que não são, totalmente, coincidentes. Ou seja, autores diferentes têm utilizado o mesmo termo (simples ou composto) com significados algo diferentes. Nesta área, não existe ainda uma terminologia estabelecida e, globalmente, aceite por todos os especialistas. **Rowena Cullen (1995)** alertou¹⁶ para a persistente inexistência de uniformização do vocabulário usado nos estudos de avaliação de bibliotecas. **Penny Garrod e Margaret Kinnell (1997)** apontaram a ausência de uma definição exata do termo *benchmarking*¹⁷ defendendo a necessidade de se encontrar vocabulário, adequado à descrição de métodos e práticas relacionados com a qualidade, que seja aceite pela maioria dos profissionais desta área. O estudo australiano de **Anne Wilson, Leanne Pitman, e Isabella Trahn (2000)**, clarifica um pouco a terminologia, dado que as

¹⁶ Rowena Cullen, *Stakeholder perceptions of university library effectiveness*, p. 438.

Citado por Anne Wilson, Leanne Pitman, e Isabella Trahn (2000).

¹⁷ Penny Garrod, Margaret Kinnell, *Benchmarking development needs in the LIS sector*, p. 115.

Citado por Anne Wilson, Leanne Pitman, e Isabella Trahn (2000).

autoras tiveram a preocupação de precisarem bem os termos que iriam utilizar ao longo do estudo, tendo em atenção definições, anteriormente, publicadas por outros autores. *Benchmarking, performance indicators e quality / best practice measures* foram os termos selecionados para aquele estudo.

A partir da década de setenta do século XX, após décadas de experiências e de reflexões¹⁸, começaram a surgir os primeiros documentos de aplicação prática e os primeiros trabalhos de investigação. Nos próximos parágrafos, tentar-se-á delinear em traços gerais um breve resumo histórico do que se passou na área da avaliação de serviços de informação desde 1974 até ao ano de 2008:

- ❖ Em Janeiro de 1974, a revista americana “Library Trends” publicava um número dedicado à avaliação de serviços de bibliotecas. Novos números relacionados com este tema seriam publicados pela mesma revista em 1981, em 1996 e em 2001. Em Setembro de 2000, a mesma revista dedicou um número à avaliação de bibliotecas digitais no qual foi publicado o famoso artigo de Tefko Saracevic.
- ❖ Em 1982, a American Library Association publicava um primeiro manual para avaliação de bibliotecas públicas da autoria de Nancy Van House e Charles McLure, entre outros.
- ❖ Em 1987, na Conferência Geral da IFLA, é apresentado um anteprojeto da obra que a UNESCO encomendou a um comitê internacional e que viria a publicar dois anos depois (1989) no âmbito do seu «Programme générale d’information et Unisist». Este documento versava, de igual modo, sobre a avaliação de bibliotecas públicas.¹⁹
- ❖ A primeira obra acerca de avaliação de bibliotecas universitárias é publicada em 1990 por Nancy Van House, Charles McClure e Beth Weil.

¹⁸ Nos EUA, já em 1932, realizavam-se relatórios das bibliotecas do Estado de Nova Iorque acerca de medições - cf. na Internet:
<<http://web.library.uiuc.edu/ahx/ala/ccard/ALAControlCard.asp?RG=29&SG=38&RS=4>>. Rowena Cullen (2001), na sua comunicação à 4th Northumbria International Conference on Performance Measurement in Libraries and Information Services refere «Public Library Association moved away from prescriptive input standards of 1933 and Minimum Standards of 1966, refocused evaluation on overall effectiveness (De Prosop, 1973)...».

¹⁹ MOORE, Nick. - **Comment mesurer l'efficacité des bibliothèques publiques.** - Paris : UNESCO, 1989. 57 p. (PGI-89/WS/3).

- ❖ Em 1995, dois departamentos da Northumbria University (Department of Information and Library Management, Department of Information Service) em parceria com a British Library (Research & Development Department) organizam conjuntamente a primeira conferência bianual acerca da medição do desempenho em bibliotecas e centros de documentação. Esta conferência tornou-se uma referência nos anos seguintes.
- ❖ No mesmo ano de 1995, a Comissão Europeia (DGXIII-E3) publica um repertório de indicadores de desempenho de bibliotecas, o qual se destinava a constituir uma obra de apoio aos quatro projetos europeus DECIDE, DECIMAL, EQLIPSE e MINSTREL.
- ❖ Em 1996, a IFLA publica as suas linhas de orientação sobre medição da qualidade / desempenho em bibliotecas universitárias. A partir desta data, os serviços de informação do ensino superior (sobretudo, as bibliotecas) passam a ter uma obra de referência para utilizarem no desenvolvimento dos seus estudos de avaliação. Nesse mesmo ano, é publicado o relatório do HEFCE (Higher Education Funding Council for England) que teve uma larga influência nas bibliotecas universitárias inglesas - baseava-se em vinte e um (21) indicadores selecionados a partir da publicação da Comissão Europeia e do projeto de norma ISO. Acabou, mesmo, por constituir um instrumento de apoio aos projetos europeus supra referidos.
- ❖ Em 1997, Peter Brophy e Peter Wynne, no âmbito do projeto e-Lib, redigem um importante relatório sobre medição de desempenho para as bibliotecas eletrónicas.
- ❖ Em 1998, surge a norma internacional – a ISO 11620 -, a qual reúne vinte e nove (29) indicadores recolhidos nas diversas obras publicadas até então.
- ❖ Em 1999, o projeto europeu EQUINOX atualiza e adapta os indicadores constantes da obra da Comissão Europeia (1995) para os aplicar na avaliação de bibliotecas eletrónicas e nos quatro projetos europeus já mencionados.
- ❖ De Maio de 2000 a Dezembro de 2001, Charles McClure, em colaboração com Wonsik "Jeff" Shim e John Carlo Bertot, conduz o projeto “E-Metrics” que envolve as vinte e quatro (24) bibliotecas da Association of Research

Libraries (ARL) fundadoras do projeto. Os responsáveis máximos pelo projeto são Sherrie Schmidt e Rush Miller. Este projeto tem por objetivo o desenvolvimento de estatísticas e de medidas de desempenho para avaliação da utilização de recursos e serviços eletrónicos nas bibliotecas especializadas das universidades.

- ❖ Em 2003 é publicada a norma ISO/TR 20983 relativa a indicadores de desempenho para serviços eletrónicos de bibliotecas²⁰, e é atualizada a norma ISO 2789, cuja 2ª edição tinha saído já em 1991, relativa a estatísticas internacionais para bibliotecas com uma nova abordagem que reflete as novas práticas das bibliotecas na época no que diz respeito a recursos eletrónicos e a automação²¹. A norma 2789 voltaria a ser revista em 2006, incluindo-se nessa 4ª edição alterações necessárias à boa aplicação da norma ISO 11620, e em 2013 com a 5ª edição.
- ❖ Mais tarde, em 2008, a ISO / TR 20983 viria a ser retirada sendo o seu conteúdo incluído na nova edição da ISO 11620²² que seria publicada nesse mesmo ano.

Ressalta desta relação que o ano de 1997 marca, com o relatório do projeto e-Lib, o início da publicação de informação científico-técnica acerca de avaliação de bibliotecas digitais. Outra ilação fácil de retirar é a constatação de que a maioria da documentação tem origem nos Estados Unidos da América, no Reino Unido e na Austrália. E, decerto, não é por acaso que são, exatamente, esses os países onde, há mais tempo, a avaliação de bibliotecas constitui uma disciplina nos cursos de biblioteconomia e de ciência da informação. Na Alemanha, talvez por influência de Roswitha Poll, a disciplina, também, começou a marcar presença nas universidades, mas, mais recentemente.

Mas, afinal, queremos avaliar... o quê? E como? Quais os indicadores mais adequados a cada tipo de serviço? Como proceder à interpretação dos dados obtidos?

²⁰ ISO/TR 20983. 2003, Information and documentation - *Performance indicators for electronic library services*. Genève : ISO.

²¹ ISO 2789. 2003, Information and documentation - *International library statistics*. Genève : ISO.

²² Está prevista uma 3ª edição da ISO 11620 para 2014.

Com efeito, nestas décadas em que a literatura científica e técnica se multiplicou, consideravelmente, divulgando inúmeros indicadores e medidas para avaliar o desempenho dos serviços de informação, o mundo em que vivemos sofreu inúmeras alterações que tiveram sérias repercussões na sociedade e nas instituições, e, conseqüentemente, nas características dos serviços de informação. Com a emergência de novas realidades, como as bibliotecas digitais, houve necessidade de criar novos métodos de avaliação ou de adaptar os existentes. Todavia, no meio de um universo tão vasto de medidas e indicadores, começa a tornar-se difícil ter uma noção exata de quantos são e ter a certeza de que se conseguem selecionar os mais adequados ao serviço a avaliar. Se nos anos trinta e quarenta do século passado, a avaliação se centrava nas coleções – a sua dimensão, os exemplares raros, etc. -, ao longo dos anos foi-se deslocando, progressivamente, para o utilizador e para o seu relacionamento com aquelas (circulação e empréstimo), acabando por se alargar a todo o serviço ou instituição. A bibliometria deu-nos uma nova visão da biblioteca: esta deixou de ser vista como um serviço fechado sobre si-mesmo e passou a ser encarada como um serviço que interage com o meio envolvente de forma dinâmica. E, desta forma, chega-se, mais recentemente, à conclusão de que não é suficiente avaliar os recursos de que o serviço dispõe (*inputs*), a satisfação do utilizador e a qualidade do serviço (*outputs*) ou até as medidas de desempenho (processos), pois, tornou-se, extremamente, importante determinar qual o impacto que o serviço de informação tem na comunidade envolvente e na sociedade. Este é, seguramente, um objetivo difícil de alcançar dada a difícil mensurabilidade dos dados a ter em conta. Todavia, houve já estudos que apontaram nesse sentido, e houve pelo menos um que foi bem-sucedido: o estudo de Rochester. Este estudo foi levado a cabo pelos bibliotecários hospitalares da área geográfica de Rochester em Nova Iorque, entre Setembro de 1990 e Março de 1991, e publicado em 1992 por J.G. Marshall da Universidade de Toronto (Faculty of Library and Information Science) no *Bulletin of the Medical Library Association* (continuado por *Journal of the Medical Library Association*). O estudo consistia num projeto de investigação que tinha por objetivo explorar o impacto dos serviços da biblioteca hospitalar na tomada de decisão clínica e que abrangeu um universo de 448 médicos, tendo o questionário sido devolvido

preenchido por 208 deles. Os resultados são impressionantes e podem ser confirmados em pormenor no resumo do artigo publicado²³. A conclusão do estudo foi a de que os médicos consideram que a biblioteca hospitalar tem um impacto significativo na tomada de decisão clínica, conclusão que fica, devidamente, reforçada pelos dados estatísticos apresentados. Só para citar alguns exemplos: em 80% dos casos, o recurso à biblioteca hospitalar modificou um qualquer aspeto da decisão clínica; em 51% dos casos, a seleção dos exames complementares de diagnóstico foi alterada; em 45% dos casos, foi alterada a seleção da terapia por fármacos; e em 19% dos casos, foi evitada a mortalidade do paciente. Com efeito, este estudo reflete bem o impacto que serviços de informação eficientes podem ter na sociedade.

No caso das bibliotecas digitais e dos arquivos digitais esse impacto afigura-se, exponencialmente, maior e muito mais difícil de identificar. Adicionalmente, os indicadores de desempenho não estão ainda bem explorados, sendo necessários mais trabalhos de investigação que elucidem quais as abordagens mais adequadas.

Historicamente, a primeira obra sobre avaliação de bibliotecas a refletir o impacto da era digital foi o documento das linhas de orientação da IFLA para bibliotecas universitárias (1996), e a primeira norma ISO a fazê-lo foi a norma 11620 publicada em 1998. Por esse motivo, analisar-se-á com maior pormenor o conteúdo de ambos os documentos. Em qualquer um deles, os indicadores selecionados requerem uma recolha de dados trabalhosa e têm, sobretudo, uma utilidade interna, pois dependem, largamente, da política de gestão de cada serviço para poderem ser utilizados em termos comparativos. Ou seja, os dados que se reportam, por exemplo, a serviços de empréstimo dependem de uma política subjacente que pode variar de serviço para serviço, já para não dizer de instituição para instituição. Ao utilizar este tipo de dados para estudos comparativos há que ser muito cauteloso estabelecendo formas de correção adequadas, evitando, assim, interpretações erróneas dos dados. A obra de referência para a avaliação de bibliotecas universitárias, publicada em 1996 pela IFLA, apresenta dezassete (17) indicadores que evidenciam aplicabilidade em

²³ Marshall JG (1992) The impact of the hospital library on clinical decision making: the Rochester study. Bull Med Libr Assoc 80(2):169-178

qualquer tipo de biblioteca universitária, independentemente, da sua dimensão ou perfil. Contrariamente ao que constava em obras anteriores, não foi menosprezado o fator tempo na avaliação – quatro indicadores implicam de forma explícita a análise da rapidez. O tempo despendido pelo utilizador num serviço de informação é de extrema importância: quanto menor for o espaço de tempo dentro do qual o utilizador encontra informação pertinente melhor, independentemente dos fatores que possam ser condicionantes. As pessoas dispõem de menos tempo para o que quer que seja. E estão, cada vez mais, habituadas a conseguirem realizar variadíssimas tarefas num curto espaço de tempo devido ao uso crescente das designadas novas tecnologias. Um serviço que não consiga prestar informações aos seus utilizadores com relativa rapidez (é evidente que há pesquisas, inevitavelmente, mais demoradas) não está preparado para o século XXI. Os indicadores seleccionados por Roswitha Poll e Peter te Boekhorst são os que seguem abaixo listados:

1. Penetração no mercado;
2. Horário de funcionamento em comparação com a procura;
3. Listagens de especialistas;
4. Utilização da coleção;
5. Utilização da coleção por assunto;
6. Documentos não utilizados;
7. Pesquisa de documentos especificados;
8. Pesquisa por assunto;
9. Velocidade de aquisição;
10. Velocidade no tratamento documental;
11. Disponibilidade;
12. Tempo de processamento do empréstimo;
13. Velocidade do Empréstimo Inter-Bibliotecas;
14. Taxa de respostas corretas no serviço de referência;
15. Acessos remotos *per capita*;
16. Satisfação do utilizador;
17. Satisfação do utilizador com os serviços de acesso remoto.

A obra acima descrita contribui com dois indicadores individualizados para a nova era da informação – a era digital. Estava-se em 1996... Volvidos dezassete anos, em 2013, as bibliotecas digitais já se encontram contempladas com indicadores específicos para a sua avaliação, os quais têm vindo a ser criados à medida que avançam e são concluídos vários projetos de investigação.

A norma publicada em 1998, e reeditada em 2008, aplica-se, genericamente, a todos os tipos de biblioteca, cabendo ao avaliador a seleção dos indicadores mais adequados de entre os vinte e nove (29) referenciados em 1998 e os quarenta e cinco (45) referenciados em 2008. Nesta lista de indicadores constavam dez (10) dos supra referidos na obra da IFLA já na primeira edição em 1998.

Os indicadores de 1998 são os seguintes:

1. Satisfação do utilizador;
2. Percentagem da população-alvo atingida;
3. Custo por utilizador;
4. Visitas à biblioteca *per capita*;
5. Custo por visita à biblioteca;
6. Disponibilidade de títulos;
7. Disponibilidade de títulos procurados;
8. Percentagem dos títulos procurados na coleção;
9. Disponibilidade alargada dos títulos procurados;
10. Utilização presencial *per capita*;
11. Taxa de utilização dos documentos;
12. Tempo médio de recuperação de documentos provenientes de depósitos;
13. Tempo médio de recuperação de documentos provenientes de áreas de livre acesso;
14. Atualização da coleção;
15. Empréstimos *per capita*;
16. Documentos emprestados *per capita*;
17. Custo por empréstimo;

18. Empréstimos por funcionário;
19. Velocidade do Empréstimo Inter-Bibliotecas;
20. Taxa de respostas corretas no serviço de referência;
21. Taxa de sucesso na pesquisa por título no catálogo;
22. Taxa de sucesso na pesquisa por assunto no catálogo;
23. Disponibilidade das instalações e do equipamento;
24. Taxa de utilização das instalações e do equipamento;
25. Taxa de ocupação de lugares sentados;
26. Disponibilidade dos sistemas automatizados;
27. Tempo médio de aquisição de documentos;
28. Tempo médio de tratamento documental;
29. Custo por título catalogado.

Quanto aos indicadores de 2008, que se seguem, estes incluem várias novidades passando de 29 a 45 no total:

1. Disponibilidade de títulos procurados;
2. Percentagem de títulos procurados na coleção;
3. Taxa de sucesso da pesquisa por assunto no catálogo;
4. Percentagem de sessões rejeitadas;
5. Exatidão da colocação na estante;
6. Tempo médio de recuperação de documentos provenientes de depósitos;
7. Velocidade do Empréstimo Inter-Bibliotecas;
8. Percentagem de Empréstimos Inter-Bibliotecas com sucesso;
9. Estações de trabalho de acesso público *per capita*;
10. Disponibilidade em horas das estações de trabalho *per capita*;
11. Área de utilizador *per capita*;
12. Lugares sentados *per capita*;
13. Número de horas de visita em comparação com a procura;
14. Funcionários (pessoal) *per capita*;

15. Volume de negócios²⁴ da coleção;
16. Empréstimos *per capita*;
17. Percentagem de depósito não usado;
18. Número de unidades de conteúdo descarregadas *per capita*;
19. Utilização presencial *per capita*;
20. Visitas à biblioteca *per capita*;
21. Percentagem de pedidos de informação submetidos eletronicamente;
22. Percentagem de utilizadores externos;
23. Percentagem do total de empréstimos efetuada a utilizadores externos;
24. Presenças do utilizador em eventos da biblioteca *per capita*;
25. Número de presenças do utilizador em sessões de formação *per capita*;
26. Taxa de ocupação de lugares sentados públicos;
27. Taxa de utilização de estações de trabalho;
28. Percentagem da população-alvo atingida;
29. Satisfação do utilizador;
30. Custo por empréstimo;
31. Custo por sessão de base de dados;
32. Custo por unidade de conteúdo descarregada;
33. Custo por visita à biblioteca;
34. Tempo médio de aquisição de documentos;
35. Tempo médio de tratamento documental (processamento de documentos);
36. Percentagem do pessoal afeto aos serviços de apoio ao utilizador;
37. Taxa de respostas corretas no serviço de referência;
38. Ratio entre as despesas com aquisições e as despesas com pessoal;
39. Produtividade dos funcionários no processamento dos documentos adquiridos;
40. Custo por utilizador;
41. Percentagem da despesa com o provisionamento de informação gasta na coleção eletrónica;
42. Percentagem do pessoal da biblioteca que fornece serviços eletrónicos;

²⁴ *Turnover* na versão em inglês.

43. Número de horas de formação por cada membro do pessoal;
44. Percentagem de recursos da biblioteca recebidos por subvenção especial ou por receitas próprias;
45. Percentagem de recursos institucionais alocados à biblioteca.

Verifica-se que se mantiveram dezassete (17) indicadores dos constantes na primeira edição da norma em 1998, tendo sido descartados doze (12) por manifesta desatualização (os relativos a pesquisa por título, disponibilidade de títulos, a número de empréstimos ou a custo por título catalogado, por exemplo), e que foram incluídos vinte e oito (28) novos indicadores. Do total de quarenta e cinco (45) indicadores, ainda (visto que passou uma década em que a evolução tecnológica foi enorme), só dez (10) dizem respeito à era digital. Na realidade, na sua maioria, aplicam-se mais a serviços híbridos do que a serviços puramente digitais. Pelo que continuamos com um problema de desadequação à realidade de muitos serviços de informação, não obstante, o facto de a norma ser mais do que satisfatória para muitos outros, os quais continuam a constituir a grande maioria dos serviços de informação. Não podemos esquecer que a evolução tecnológica no mundo, e até dentro de uma mesma comunidade como a União Europeia, se faz a várias velocidades e que as necessidades de serviços de informação de países desenvolvidos com acesso a financiamentos estatais ou internacionais são bem diferentes das de serviços de informação localizados em áreas do planeta bem menos bafejadas por apoios financeiros e por infraestruturas avançadas nas redes de informação e de comunicação. A velocidade da transmissão de dados beneficia ou penaliza consoante o ponto do globo em que nos encontramos, e nem mesmo as instalações de unidades de investigação ou de instituições de ensino superior escapam a essa realidade. Segundo o relatório da Akamai “The State of the Internet”²⁵, relativo ao primeiro trimestre de 2013, publicado no *website* da empresa em 23 de Agosto de 2013, enquanto a média global de velocidade média de acesso à Internet se situou em 3,1 mbps²⁶, o Brasil e a China apresentaram uma velocidade média de, apenas, 2,19

²⁵ Akamai - “The State of the Internet : 1st quarter 2013 report”:
http://www.akamai.com/stateoftheinternet/?WT.ac=soti_banner

²⁶ mbps = megabit por segundo

mbps, e Portugal de 4,99 mbps (abaixo do seu valor no segundo trimestre de 2011 que era de 5,37 mbps) ficando no 34º lugar da lista a nível mundial. A Coreia do Sul foi o país que apresentou o melhor desempenho com o valor de 14,18 mbps, pelo que lidera a lista. Em segundo lugar ficou o Japão com 10,23 mbps, em terceiro Hong Kong com o valor de 10,73 mbps, em quarto a Suíça com 10,09 mbps, e em quinto a Holanda com 9,52 mbps. Os Estados Unidos da América (EUA) tiveram um valor de 8,39 mbps ficando em nono lugar (e figurando, ainda, nos dez primeiros da lista), o Reino Unido ficou pelos 7,59 mbps e a Alemanha pelos 6,87 mbps, conforme se pode verificar no gráfico da figura 1 abaixo reproduzido.

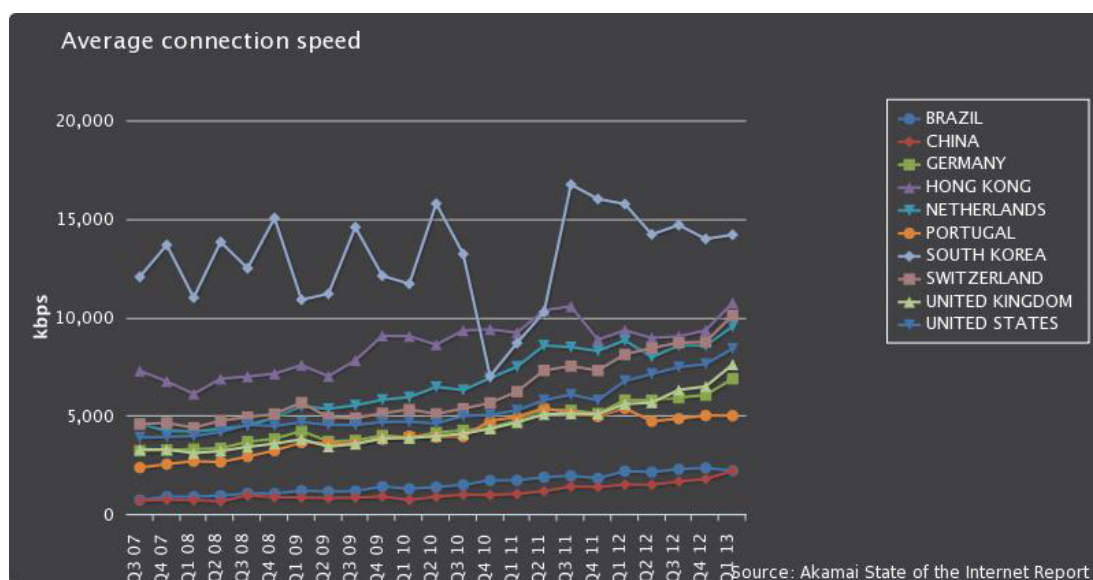


Figura 1 – Gráfico comparativo de dados de dez países (Brasil, China, Alemanha, Hong Kong, Holanda, Portugal, Coreia do Sul, Suíça, Reino Unido e Estados Unidos da América) quanto à velocidade média de acesso à Internet no primeiro trimestre de 2013.²⁷

Quer o Reino Unido quer a Alemanha só se destacam de Portugal a partir do primeiro trimestre de 2012 após a introdução das medidas do *Programa de Ajustamento Económico para Portugal*, nas suas primeira (Verão de 2011)²⁸ e segunda revisões (Outono/Inverno de 2011)²⁹. No segundo trimestre de 2011, enquanto Portugal se situava nos 5,37 mbps - o seu melhor valor desde sempre -, o

²⁷ Gráfico gerado a partir do website da Akamai na visualização de dados do relatório supra-referido: http://www.akamai.com/stateoftheinternet/?WT.ac=soti_banner

²⁸ http://ec.europa.eu/economy_finance/publications/occasional_paper/2011/pdf/ocp83_en.pdf

²⁹ http://ec.europa.eu/economy_finance/publications/occasional_paper/2011/pdf/ocp89_en.pdf

Reino Unido situava-se, ainda, nos 5,06 mbps e a Alemanha nos 5,29 mbps – igualmente, os seus melhores valores desde sempre. O gráfico da figura 2 permite visualizar estes dados de forma bem mais evidente do que o gráfico da figura 1 que contém, maior número de dados.

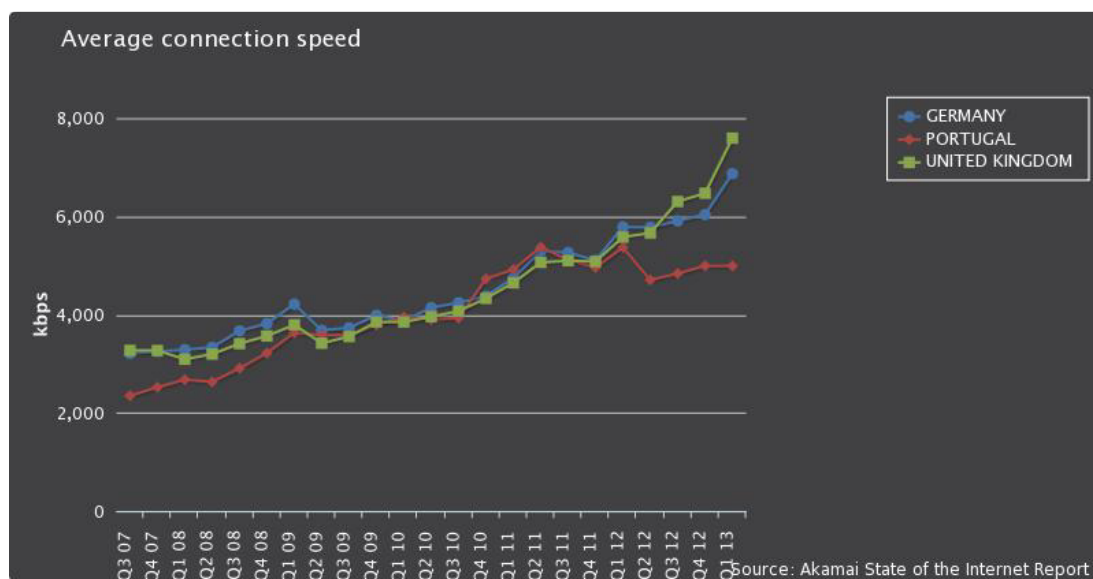


Figura 2 - Gráfico comparativo de dados de três países da União Europeia (Alemanha, Portugal e Reino Unido) quanto à velocidade média de acesso à Internet no primeiro trimestre de 2013³⁰.

Já quanto à velocidade média de picos de acesso à Internet constantes da figura 3, o valor global médio foi de 18,4 mbps e Portugal situou-se nos 33,67 mbps tendo ficado em 13º lugar a nível mundial. A China ficou pelos 9,47 mbps e o Brasil pelos 18,31 mbps, ligeiramente abaixo da média global. No topo da lista surge Hong Kong em primeiro lugar com 65,73 mbps, seguindo-se o Japão com 47,37 mbps em segundo lugar, e ficando a Coreia do Sul em quarto lugar (ultrapassada pela Roménia) com 44,93 mbps. A Suíça surge no sétimo lugar da lista tendo apresentado valores na ordem dos 39,88 mbps, e a Holanda no nono lugar com 36,48 mbps. Abaixo dos primeiros dez no topo da lista, figuram os EUA com 35,88 mbps, o Reino Unido com valores na ordem dos 35,26 mbps, Portugal com o valor acima referido, a Alemanha com 31,24 mbps, e o Brasil e a China com os valores,

³⁰ Gráfico gerado a partir do *website* da Akamai na visualização de dados do relatório supra-referido: http://www.akamai.com/stateoftheinternet/?WT.ac=soti_banner

igualmente, acima mencionados. Os dados relativos a Portugal podem ser visualizados no gráfico da figura 4 com maior pormenor.

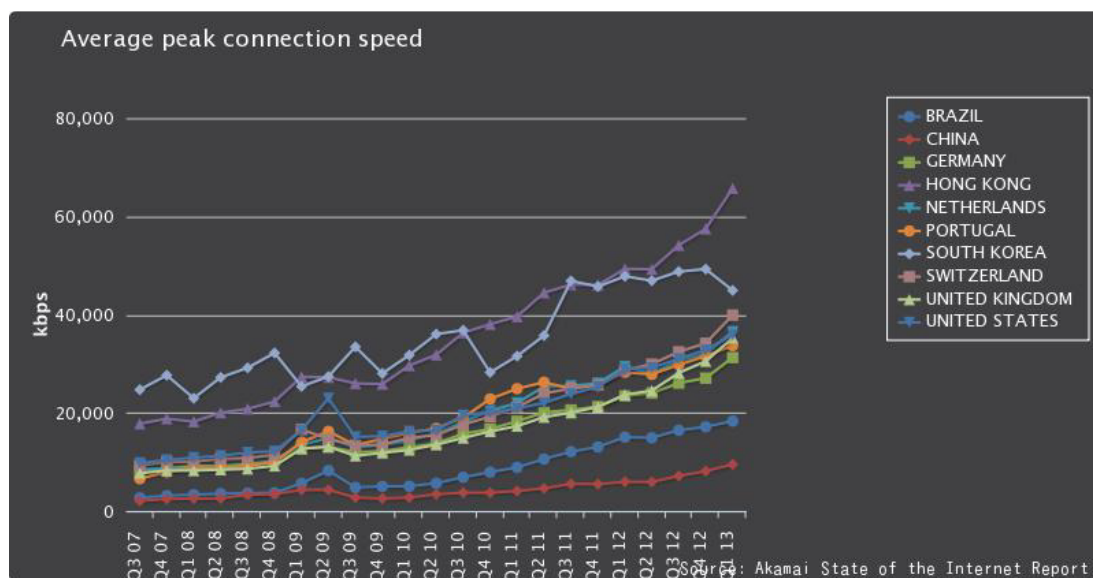


Figura 3 - Gráfico comparativo de dados de dez países (Brasil, China, Alemanha, Hong Kong, Holanda, Portugal, Coreia do Sul, Suíça, Reino Unido e Estados Unidos da América) quanto à velocidade média de picos de acesso à Internet no primeiro trimestre de 2013³¹.

Tal como na lista da velocidade média de acesso à Internet, verifica-se que Portugal inicia uma queda nos valores no terceiro trimestre de 2011, sendo já ultrapassado pela Holanda neste mesmo trimestre, pelos EUA no primeiro trimestre de 2012 e pelo Reino Unido no primeiro trimestre de 2013 quanto à velocidade média de picos de acesso à Internet. No segundo trimestre de 2011, Portugal apresentava o seu melhor valor desde sempre com 26,20 mbps, a Holanda com 25,33 mbps, os EUA com 22,00 mbps, a Alemanha com 20,08 mbps e o Reino Unido com 19,11 mbps. Note-se que a curva da generalidade dos países tem sido ascendente, pelo que os outros três países apresentavam, também, os seus melhores valores desde sempre. No gráfico da figura 1, o único país que apresenta uma curva irregular é, curiosamente, a Coreia do Sul. Hong Kong, também, oscila um pouco os valores de trimestre para trimestre, mas, com menos irregularidade. Nos países europeus e americanos, e na China de forma mais suave, é evidente uma queda generalizada dos valores entre o primeiro e o terceiro trimestres de 2009 e uma retoma da curva

³¹ Idem.
46

ascendente, logo, no quarto trimestre de 2009 (a China inicia a queda no quarto trimestre de 2008 e continua a cair nos dois trimestres seguintes embora de forma mais suave), tendo os valores do primeiro trimestre de 2009 sido ultrapassados cerca dois trimestres a um ano depois na Europa e cerca de um ano depois nos EUA e na China. Esta queda generalizada ter-se-á devido, decerto, à crise económica mundial de 2008. Por comparação, percebemos que a quebra na linha correspondente a Portugal (nos gráficos das figuras 3 e 4) no terceiro trimestre de 2011 é recuperada, dois trimestres depois, no primeiro trimestre de 2012, enquanto a quebra da linha correspondente a Portugal (nos gráficos das figuras 1 e 2) ainda não foi recuperada no primeiro trimestre de 2013. Há, de novo, uma curva, tendencialmente, ascendente, mas, que ascende a um ritmo mais lento perdendo posições no ranking internacional.



Figura 4 - Gráfico comparativo de dados de três países da União Europeia (Alemanha, Portugal e Reino Unido) quanto à velocidade média de picos de acesso à Internet no primeiro trimestre de 2013³².

Cerca de dois terços da população mundial nem sequer tem acesso à Internet. Mark Zuckerberg³³, com o seu projeto “internet.org”³⁴, pretende dar acesso gratuito à

³² Ibidem.

³³ Mark Zuckerberg em entrevista à CNN - vídeo publicado a 21 Agosto de 2013: <http://edition.cnn.com/video/data/2.0/video/business/2013/08/21/newday-cuomo-zuckerberg-internet-access.cnn.html> ; <http://www.youtube.com/watch?v=F8N3wGjiPYg>

Internet aos cerca de 5 biliões de pessoas que, neste momento, ficam excluídas em todo o mundo³⁵, através do Facebook. O projeto de Zuckerberg, não obstante os interesses comerciais que se lhe possam associar (é um projeto conjunto de sete empresas: Facebook, Ericsson, MediaTek, Nokia, Opera, Qualcomm e Samsung), é o de conseguir o reconhecimento do direito à conectividade, isto é, o direito de acesso à Internet como um direito humano³⁶. O direito à informação é reconhecido na *Constituição da República Portuguesa*³⁷ desde 1976, bem como o direito de acesso aos dados informáticos que lhe digam respeito por parte de cada cidadão e o direito de acesso aos arquivos e registos administrativos (Art.º 268º) – inseridos em revisões constitucionais posteriores -, pelo que o direito de acesso à Internet acaba por se encontrar implícito. Todavia, o reconhecimento do direito de acesso à Internet numa sociedade da informação e, cada vez mais, do conhecimento que faz passar pela Internet boa parte do cumprimento de deveres e obrigações cívicas, e o próprio direito de acesso à informação, pode vir a diminuir as desigualdades e ajudar a reduzir a exclusão informacional em todo o mundo. Por enquanto, este projeto não apresenta soluções para aqueles que não têm acesso a um computador ou a um telemóvel, pelo que uma parte da população mundial continuará a ficar excluída. Por seu lado, a Google tenta melhorar a conectividade em zonas geográficas em que a captação de sinal da Internet é difícil, usando balões para retransmitir o sinal da Internet. Em Junho de 2013³⁸, foram levados a cabo os primeiros lançamentos de balões, na Nova Zelândia, no âmbito deste projeto apelidado de “Project Loon”³⁹. O objetivo é, igualmente, o de alargar o acesso à Internet a todos no mundo inteiro.

³⁴ <http://www.internet.org/>

³⁵ O artigo que saiu no “The New York Times” a 20 de Agosto de 2013:

<http://www.nytimes.com/2013/08/21/technology/facebook-leads-an-effort-to-lower-barriers-to-internet-access.html?hp>

³⁶ Notícia no jornal “Público” acerca do projeto “internet.org”:

<http://www.publico.pt/tecnologia/noticia/zuckerberg-reune-gigantes-da-tecnologia-para-elevar-acesso-a-internet-a-direito-humano-1603595>

Artigo de Mark Zuckerberg sobre o projeto “internet.org”:

<https://www.facebook.com/isconnectivityahumanright/isconnectivityahumanright.pdf>

³⁷ Constituição da República Portuguesa em vigor – revisão constitucional de 2005:

<http://www.parlamento.pt/Legislacao/Documents/constpt2005.pdf>

³⁸ Notícia no jornal “Público” acerca do projeto “Project Loon”:

<http://www.publico.pt/tecnologia/noticia/google-usa-baloes-gigantes-para-testar-acesso-a-internet-1597427#/0>

³⁹ Loon for all – Balloon-powered Internet for all <http://www.google.com/loon/>

Interesses comerciais à parte, temos em marcha uma verdadeira intenção de democratização da Internet e do acesso à informação.

O principal motor da (r)evolução tecnológica para a era digital foi o financiamento através de programas de investigação e desenvolvimento (I&D) que abraçaram inúmeros projetos a nível internacional. E os sistemas de informação digitais, nomeadamente, os arquivos e bibliotecas digitais, são, em boa parte, fruto desses programas de I&D.

Muitos têm sido os projetos e programas de investigação desenvolvidos nos últimos anos nesta área, tornando-se difícil evocá-los a todos sem que se proceda a uma inventariação prévia. Em anexo, listam-se vários projetos de avaliação de serviços de informação identificando-os pela sua sigla e/ou designação por extenso, indicando as datas de início e de conclusão, determinando a sua origem geográfica e, sempre que possível, apresentando uma breve descrição dos mesmos⁴⁰.

De entre os projetos de investigação que tiveram importância para a avaliação de sistemas de informação digitais destacamos os seguintes:

- ❖ Projetos de avaliação de sistemas de informação digitais:
 - o projeto europeu EQUINOX;
 - o projeto inglês MIEL;
 - o projeto americano E-Metrics;
 - o projeto americano DigiQUAL;
 - o projeto europeu DELOS⁴¹;

Como é possível verificar, os três projetos pioneiros em avaliação de bibliotecas digitais são o EQUINOX, o MIEL e o E-Metrics, pelo que se apresenta uma breve descrição dos mesmos abaixo.

EQUINOX (1998-2000)⁴² - Library Performance Measurement and Quality Management System era um projeto europeu que se desenvolveu com base nas

⁴⁰ Cf. a tabela 1 no anexo 1.

⁴¹ Acessível na Internet: «<http://www.delos.info/>».

conclusões dos projetos europeus DECIDE, EQLIPSE, DECIMAL, MINSTREL e CAMILE, tendo tomado a liderança a nível mundial por ter sido o primeiro projeto a desenvolver indicadores de desempenho para bibliotecas digitais. Este projeto foi levado a cabo por um consórcio de sete parceiros europeus e coordenado por Peter Brophy do CERLIM sediado na Manchester Metropolitan University. Os parceiros eram os seguintes: Manchester Metropolitan University, UK; National Microelectronics Application Centre, Ltd, UK; Fretwell-Downing Informatics Ltd., UK; Dublin City University Library, Ireland; Universitäts und Landesbibliothek Münster, Germany; Universitat Oberta de Catalunya, Spain; University of Stockholm Library, Sweden. O principal objetivo do projeto era o de desenvolver indicadores de desempenho para bibliotecas digitais, tal como já foi referido. Os objetivos específicos eram os seguintes:

- desenvolver um software integrado de apoio aos gestores de bibliotecas europeus na gestão de serviços progressivamente híbridos (tradicionais e eletrónicos) de forma eficaz e eficiente;
- desenvolver um conjunto normalizado de indicadores de desempenho adequado à “biblioteca híbrida” e procurar consenso internacional nesta área;
- identificar os dados a recolher para desenvolver esses indicadores de desempenho;
- fornecer software que encoraje todos os gestores de bibliotecas a introduzirem um nível adequado de gestão da qualidade, eliminando as restrições da ISO 9002;
- validar e testar o protótipo do sistema numa amostra de bibliotecas;
- levar a cabo uma série de demonstrações, em larga escala, em bibliotecas de toda a Europa;
- divulgar o modelo por toda a Europa;
- assegurar que a Europa mantém a liderança internacional nesta área.

Os indicadores desenvolvidos pelo projeto são catorze e listam-se nas linhas seguintes:

⁴² Acessível na Internet: <<http://equinox.dcu.ie/>>. O texto que se segue constitui uma adaptação livre de parte do conteúdo do *website*.

1. Percentagem da população atingida pelos serviços da biblioteca eletrónica;
2. Número de sessões em cada serviço da biblioteca eletrónica por membro do público-alvo;
3. Número de sessões remotas nos serviços da biblioteca eletrónica por membro do universo de utilizadores;
4. Número de documentos e entradas (registos) visualizados por sessão para cada serviço da biblioteca eletrónica;
5. Custo por sessão para cada serviço da biblioteca eletrónica;
6. Custo por documento ou entrada (registo) visualizado para cada serviço da biblioteca eletrónica;
7. Percentagem de pedidos de informação submetidos eletronicamente;
8. Taxa de utilização dos computadores disponíveis na biblioteca;
9. Número de horas em que os computadores da biblioteca estão disponíveis para consulta por membro do universo de utilizadores;
10. Sessões rejeitadas como percentagem do número total de tentativa de início de sessão;
11. Percentagem do total da verba destinada a aquisições gasta na aquisição de serviços eletrónicos;
12. Número de inscrições por membro do universo de utilizadores em cursos de formação na utilização de serviços da biblioteca eletrónica;
13. Percentagem de pessoal da biblioteca que desenvolve, gere e fornece serviços da biblioteca eletrónica e que forma utilizadores na utilização dos mesmos;
14. Satisfação do utilizador com os serviços da biblioteca eletrónica.

As conclusões do projeto indicam que o mercado potencial para este tipo de sistema se alarga a todo o tipo de biblioteca a nível mundial. Mas, a sua imediata aplicabilidade restringir-se-á a bibliotecas com sistemas de tecnologia da informação já implantados. A expansão do mercado de tecnologias da informação para bibliotecas far-se-á, previsivelmente, mediante a adesão das bibliotecas existentes a sistemas de tecnologia da informação e da comunicação, e não tanto através do

aparecimento de novas bibliotecas.⁴³ Chega-se, igualmente, à conclusão de que, possivelmente, as bibliotecas poderão preferir proceder à aquisição do sistema EQUINOX como um complemento para o principal sistema de informação que utilizam.

MIEL (1998)⁴⁴ - Management Information for the Electronic Library

Trata-se de um estudo de menor dimensão, integrado nas iniciativas e atividades FIGIT,⁴⁵ que teve por objetivo perspetivar uma investigação detalhada acerca das necessidades dos gestores de bibliotecas em termos de gestão de informação e de medição do desempenho na era das bibliotecas eletrónicas. Apresentam-se sugestões para o desenvolvimento de um conjunto de medidas de desempenho para a biblioteca eletrónica, e aposta-se na cooperação em investigação com investigadores americanos a trabalharem na mesma área, como é o caso de Charles McClure (Syracuse University). Pretende-se, ainda, elaborar um manual de técnicas que permita aos gestores de bibliotecas reunir, analisar, e utilizar técnicas de gestão da informação e de medição de desempenho adequadas às bibliotecas eletrónicas.

E-Metrics (2000-2001)

Projeto desenvolvido pela Association of Research Libraries (ARL) nos EUA, de Maio de 2000 a Dezembro de 2001, por iniciativa de vinte e quatro (24) das suas bibliotecas associadas. Charles McClure, Wonsik "Jeff" Shim e John Carlo Bertot, coordenaram o projeto. Os responsáveis máximos pelo projeto são Sherrie Schmidt e Rush Miller, os quais coordenaram o projeto dentro da ARL. O projeto foi contratado com o Florida State University's Information Use Management and Policy Institute. Este projeto tem por objetivo o desenvolvimento de estatísticas e de medidas de desempenho para avaliação da utilização de recursos e serviços eletrónicos nas bibliotecas especializadas das universidades.

As medidas propostas no âmbito do projeto e divulgadas na fase 3 são as seguintes:

⁴³ A nível mundial, esta afirmação estará correta, todavia, a um nível geográfico mais restrito poderá não ser tanto assim... Repare-se no caso de Portugal: a criação da Rede de Leitura Pública alterou por completo o mercado de sistemas de informação para serviços de informação em território português.

⁴⁴ Acessível na Internet: <<http://www.ukoln.ac.uk/dlis/models/studies/mis/mis.rtf>>.

⁴⁵ Follett Implementation Group on Information Technology (UK).

1. Recursos eletrónicos pagos a editores
 - R1 Número de revistas eletrónicas em texto integral
 - R2 Número de obras de referência eletrónicas
 - R3 Número de e-books
2. Utilização de recursos e serviços eletrónicos em rede
 - U1 Número de transações eletrónicas de referência
 - U2 Número de logins (sessões) em bases de dados eletrónicas
 - U3 Número de pesquisas em bases de dados eletrónicas
 - U4 Itens pedidos em bases de dados eletrónicas
 - U5 Visitas virtuais ao sítio na Internet e ao catálogo da biblioteca
3. Despesas com recursos eletrónicos e infraestrutura relacionada
 - C1 Custo de revistas eletrónicas em texto integral
 - C2 Custo de obras de referência eletrónicas
 - C3 Custo de e-books
 - C4 Despesas da biblioteca em artigos bibliográficos, redes de dados e consórcios
 - C5 Despesas externas com artigos bibliográficos, redes de dados e consórcios
4. Estatísticas relacionadas com atividades de digitalização da biblioteca
 - D1 Tamanho da coleção digital da biblioteca
 - D2 Uso da coleção digital da biblioteca
 - D3 Custo da construção e gestão da coleção digital da biblioteca
5. Indicadores de desempenho sugeridos
 - P1 Percentagem de transações eletrónicas de referência de entre o número total de transações na área de referência
 - P2 Percentagem de visitas virtuais à biblioteca de entre o número total de visitas
 - P3 Percentagem de e-books de entre todas as monografias

Após a conclusão do projeto, a investigação continuou apoiando-se no conjunto de dados divulgados em 29 de Julho de 2002.

Sherrie Schmidt e Rush Miller continuam como coordenadores e Martha Kyrillidou é o contacto para eventuais participantes.

DigiQUAL (2003-) - Coordenado por Martha Kyrillidou e Sara Giersch, é o sucessor do projeto SERVQUAL e parceiro do projeto LibQUAL+. Continua a trabalhar no desenvolvimento de métricas para o mundo digital em 2013⁴⁶.

DELOS (2004-2006) – A Network of Excellence on Digital Libraries.

O *cluster* WP7 deste projeto é coordenado por Norbert Fuhr e aborda a avaliação de bibliotecas digitais. Há vários estudos publicados, podendo considerar-se que um artigo publicado na Springer em 2007, sob o título “Evaluating Digital Libraries”⁴⁷, em coautoria pelos membros da equipa, seja o mais relevante. O relatório do projeto foi publicado no *website* em 2006⁴⁸. O projeto foi continuado por um outro projeto que teve início em 2008, tendo o *website* inicial sido “congelado” em 2009 e sido reencaminhado para o novo *website* da DL.org⁴⁹.

DL.org (2008-) – Digital Library Interoperability, Best Practices and Modelling Foundations

Projeto que deu continuidade ao projeto europeu DELOS. Inclui seis grupos de trabalho: Content; Functionality; Users; Policy; Quality; Architecture. E tem como peritos nomes como o de Tefko Saracevic e Nicola Ferro na área da qualidade de bibliotecas digitais. Na página da comunidade são delineados o objeto e o objetivo da DL.org como se pode ler no excerto que se segue:

«The core activities and main outputs of DL.org have mainly targeted the Library and Information Science (LIS) Community, spanning educationalists, students, practitioners, and decision makers. DL.org has also engaged closely with the computer science and research community to ensure a focused approach to future advancement in the digital library field. To underpin and shape outputs, DL.org has harnessed global expertise through six Working Groups (Content, Functionality, User, Policy, Quality and Architecture), a Liaison Group and an External Advisory

⁴⁶ <http://www.digiqua.org/home>

⁴⁷ FUHR, Norbert, et al. – Evaluation of digital libraries. *International Journal on Digital Libraries*. 8:21-38 (2007). DOI 10.1007/s00799-007-0011-z

⁴⁸ http://www.delos.info/index.php?option=com_content&task=view&id=28&Itemid=262

⁴⁹ <http://www.dlorg.eu/>

Board. Together, all these professionals make up the DL.org stakeholder community on multiple levels: developing or enhancing digital libraries, making these systems interoperable, learning and teaching, as well as assessing the extent to which a specific digital library conforms with the DL.org Digital Library Reference Model.»⁵⁰

Quanto às abordagens de investigação mais utilizadas na avaliação de bibliotecas digitais são poucos os estudos que se lhes dedicam.

Covey (2002), em representação da Digital Library Federation (DLF), conduziu uma série de entrevistas, entre Novembro de 2000 e Fevereiro de 2001, em 24 das 26 instituições que são membros da DLF para coligir dados acerca das experiências de avaliação da usabilidade e utilização das coleções digitais das principais bibliotecas digitais que existiam à data da última viragem de século e de milénio. A taxa de resposta rondou os 84% dentro das 24 instituições participantes. Os entrevistados identificaram seis desafios em termos de avaliação: recolher dados que sejam significativos e direcionados para o objetivo dos estudos; desenvolver competências para a recolha, análise e tratamento, interpretação, apresentação e aplicação de dados; elaborar planos de avaliação globais; planejar a avaliação como uma atividade fundamental; compilar e gerir os dados recolhidos para efetuar a avaliação; coligir informação suficiente que permita compreender as tendências de utilização da biblioteca pelos seus utilizadores. A generalidade dos entrevistados aponta os estudos do utilizador como a principal abordagem de avaliação aplicada, o que reflete, inequivocamente, uma filosofia de gestão de serviços centrada no utilizador. Os estudos são conduzidos das mais variadas formas: em alguns casos há pessoal especialmente dedicado a esta função, noutros é dada formação a pessoal para que a possa desempenhar, noutros, ainda, são contratadas pessoas ou empresas em regime de *outsourcing* para o efeito. Todas estas bibliotecas usam múltiplos métodos de investigação para a realização dos estudos de utilizador, sendo consensual entre os entrevistados que essa é uma condição necessária para uma completa compreensão e confirmação e/ou cruzamento dos resultados. Todavia, isto coloca questões importantes em termos de alocação e gestão de recursos para a realização destes

⁵⁰ <http://www.dlorg.eu/index.php/strategy-achievements>

estudos que são complexos e demorados, tornando-se facilmente dispendiosos, e cuja interpretação de resultados e aplicação das ilações daí resultantes nem sempre é bem-sucedida devido a dificuldades na organização do processo. Os estudos de utilizador levados a cabo com sucesso, mais frequentemente, nestas bibliotecas da DLF são os questionários, os grupos de foco, e os protocolos de utilizador. Os questionários foram utilizados para efetuar a avaliação de padrões de utilização, frequência de utilização, facilidade de utilização e sucesso de utilização dos serviços; de necessidades dos utilizadores, expectativas dos utilizadores, perspetivas dos utilizadores, prioridades e preferências na utilização das coleções e dos serviços; satisfação do utilizador com produtos externos, coleções, serviços, pessoal e sítios web; a qualidade dos serviços; mudanças de atitude e de opinião do utilizador; relevância dos serviços ou das coleções para o currículo. No que diz respeito aos recursos eletrónicos, a abordagem genericamente utilizada é a *transaction log analysis*. Note-se que, neste relatório, os testes de usabilidade não são referidos entre as abordagens principais. Este estudo, não obstante o seu título, retrata a realidade existente no início do século XXI centrando a sua abordagem, ainda, nas bibliotecas híbridas e não, exclusivamente, nas bibliotecas digitais quer estejam ou não associadas a bibliotecas analógicas. Tem a vantagem de, ao efetuar um levantamento da avaliação de serviços de informação de natureza híbrida, permitir equacionar os múltiplos desafios que enfrentam as bibliotecas nesta mudança de plataforma dos sistemas de informação, revelando uma preocupação geral dos entrevistados com as tendências de utilização das bibliotecas pelos utilizadores.

Jansen (2009), ao tentar compreender a interação entre o utilizador e a *World Wide Web* através da *web log analysis*, interroga-se sobre a possibilidade de conseguir desenhar o perfil de um utilizador que pesquisa na Internet, e até que ponto o conseguimos fazer, partindo dos dados recolhidos em cada sessão de pesquisa, mesmo quando o início de sessão não se efetuou por autenticação. A resposta é afirmativa, e evidencia que o conseguimos fazer até um nível de particularidade impressionante, o que levanta várias questões em termos de privacidade, e tem

implicações na área de desenvolvimento de sistemas, de publicidade na Internet, e quanto a políticas de utilização deste tipo de *profiling*.

Concluindo, revela-se ser necessário desenvolver estudos orientados para os utilizadores (testes de usabilidade, inquérito, análise dos percursos seguidos pelos utilizadores em cada sessão) que identifiquem as suas reais necessidades de informação e o seu comportamento informacional, não obstante o facto de a obtenção de uma amostra populacional representativa poder constituir um problema no âmbito do trabalho de investigação. A partir do resultado destes estudos poderá ser possível extrair conclusões que nos conduzam à elaboração de modelos de avaliação de SI digitais mais adequados e mais eficientes e direcionados para áreas ainda pouco exploradas como a presença e a importância dos sistemas de organização do conhecimento na recuperação da informação em arquivos e bibliotecas digitais.

1.2. A usabilidade em arquivos e bibliotecas digitais – estado da arte

Os estudos de usabilidade em SI do tipo bibliotecas digitais são já em algum número e alguns deles reportam-se a projetos com alguma ou bastante notoriedade. Já quanto a estudos de usabilidade em arquivos digitais a informação recuperada era inexistente até 2008. O modelo resultante deste projeto de investigação foi elaborado, no início de 2010 (o trabalho laboratorial deveria ter decorrido em 2010 como previsto no plano de doutoramento) tendo por objeto não só as bibliotecas digitais, mas, também, os arquivos digitais. O trabalho laboratorial foi realizado na Universidade de Aveiro, em Maio de 2011, e incidiu sobre a avaliação quer de arquivos digitais quer de bibliotecas digitais. É, apenas, em 2013 que surge um modelo formal desenhado, especificamente, para arquivos digitais e que se encontra descrito num artigo da autoria de Nicola Ferro e Gianmaria Silvello da Universidade de Padua publicado na

“Information Processing and Management”, pela Elsevier, em Julho deste ano.⁵¹ Mas, este não é um modelo dedicada à avaliação de arquivos digitais e sim um modelo para a criação de arquivos digitais. Inspirando-se no modelo 5S das bibliotecas digitais, Ferro e Silvello criaram o primeiro modelo formal, hierárquico, para o desenvolvimento de arquivos digitais.

Os arquivos e as bibliotecas digitais são SI complexos, cuja estrutura, conteúdo e serviços podem variar ao longo do tempo. Uma pequena alteração em qualquer uma destas componentes pode modificar o desempenho do SI suscitando um processo de avaliação do mesmo. Boa parte dos estudos de usabilidade que se realizam são aplicados de forma reativa e não como parte de um processo regular de avaliação e controlo do SI. Não obstante, foram já desenvolvidos vários modelos e técnicas de usabilidade aplicáveis em situações e casos específicos que podem e devem ser aplicados em casos análogos. Dada a enorme variedade de vertentes a avaliar nestes SI digitais, os modelos conhecidos não cobrem todos os aspetos a avaliar num SI deste tipo, pelo que se torna necessária a criação de um novo modelo que permita avaliar os SI pela perspetiva do utilizador, nomeadamente, durante o acesso aos sistemas de recuperação da informação e o processo de recuperação da informação. Os estudos de usabilidade em bibliotecas digitais abordam a interação homem-máquina e a facilidade de acesso à interface do utilizador, bem como a facilidade de recuperar informação que satisfaça as necessidades de informação do utilizador a partir daquela, para além da opinião do utilizador acerca das ferramentas e funcionalidades de pesquisa disponíveis na biblioteca digital.⁵²

Nos últimos vinte anos foram vários os estudos de usabilidade de bibliotecas digitais levados a cabo e foram desenvolvidos alguns modelos e técnicas, bem como foram identificados critérios ou fatores, de avaliação da usabilidade de bibliotecas digitais. A interface do utilizador tem sido objeto de estudo com frequência, neste tipo de

⁵¹ Cf. Ferro (2013).

⁵² Cf. Chowdhury (2011), p.153.

estudos, pelo facto de ser um elemento central de qualquer SI, nomeadamente, na interação homem-máquina, e, conseqüentemente, da própria ciência da informação.

A maioria dos estudos de usabilidade de bibliotecas digitais tem avaliado a eficácia e a eficiência da interface do utilizador e a satisfação do utilizador, centrando-se em dois tipos de técnicas – as empíricas e as analíticas⁵³. As primeiras implicam a realização de testes aos SI por intermédio de utilizadores e as segundas efetuam a avaliação dos SI pelo próprio perito através de teorias e métodos comprovados.

Podemos enunciar treze (13) **técnicas utilizadas em estudos de avaliação da usabilidade**:

1. Testes de usabilidade (onde podemos incluir os testes de usabilidade com *eyetracking*);
2. Questionários;
3. Entrevistas;
4. Grupos de foco;
5. *Think-aloud* ou “pensar em voz alta”;
6. *Transaction/Web Log Analysis* ou “análise dos percursos seguidos pelos utilizadores em cada sessão”;
7. *Cognitive walkthrough* ou “passo-a-passo cognitivo”;
8. Inspeção da usabilidade;
9. *Card sort*;
10. Avaliação heurística;
11. Análise de reclamações;
12. Prototipagem de papel;
13. CASSM (*Concept-Based Analysis of Surface and Structural Misfits*) ou análise baseada em conceitos de desajustes superficiais e estruturais.

Popp (2001) levou a cabo um estudo no âmbito da Association of Research Libraries (ARL) cujos resultados revelaram que a maioria das bibliotecas utilizava uma

⁵³ Cf. Marchionini, Komlodi (1998) e Choudhury, Hobbs, Lorie (2002).

combinação de métodos nos seus estudos de utilizadores sendo o “Passo-a-passo cognitivo” o mais usado com cerca de 52% das ocorrências, seguindo-se a avaliação individual do pessoal com 50%, os grupos de foco com 44%, os questionários *online* com 40%, e os testes de usabilidade e a *Log Analysis* com 37%.

Os **testes de usabilidade** são a técnica em que a informação obtida está mais ligada ao utilizador e em que a mesma está mais pormenorizada. Os participantes num teste de usabilidade são utilizadores reais com tarefas reais cujo comportamento e expressão verbal e não-verbal é observada e registada e/ou gravada. O objetivo do estudo é sempre o de melhorar a usabilidade de um produto ou serviço e os dados recolhidos são analisados permitindo o diagnóstico de eventuais problemas de usabilidade e a elaboração de recomendações no sentido de corrigir os problemas encontrados⁵⁴. Quando realizados com *eyetrackers* designam-se de testes de usabilidade com *eyetracking*, permitindo a análise dos movimentos oculares dos participantes no monitor do *eyetracker* aquando da visualização do SI a ser testado.

Os **questionários** que se incluem no método de investigação por inquérito, tal como as entrevistas, permitem a recolha de uma grande quantidade de dados de uma forma pouco dispendiosa, pelo que são bastante usados, sobretudo nas ciências sociais e humanas, abarcando amostras populacionais que podem ser bastante alargadas, sobretudo, desde que são realizados em plataformas digitais.

As **entrevistas** que se incluem, igualmente, no inquérito, permitem recolher informação de um modo semiformal e identificar as expectativas, conceitos, pensamentos, necessidades, experiências do entrevistado. Porém, não permite ver com exatidão a forma como o participante/entrevistado se comporta em interação com o produto, mas sim que racionalização aquele faz do seu comportamento mesmo que o faça inconscientemente. A análise é realizada pelo participante e não pelo observador e é assíncrona e não síncrona como no caso do *think-aloud* e dos testes de usabilidade.

⁵⁴ Cf. J.F. Dumas, J.C. Redish (1993).

Os **grupos de foco** incluem-se, igualmente, no inquérito. O inquérito por entrevista de grupos de foco consiste na realização de uma entrevista a um grupo de participantes e no âmbito dos estudos de usabilidade é usada de múltiplas formas. O ponto fraco desta técnica é o mesmo já enunciado na técnica anterior - não permite ver com exatidão a forma como os participantes/entrevistados se comportam em interação com o produto, mas sim que racionalização fazem do seu comportamento mesmo que o façam inconscientemente, com a agravante de poderem ser influenciados pelos outros participantes que constituem o grupo. A análise é realizada pelo participante e não pelo observador e é assíncrona.

O *think-aloud*, ou “pensar em voz alta”, permite captar de forma síncrona as reações do participante no estudo, sejam elas de irritação, de hesitação, ou de confusão, entre outras, bem como permite que o observador identifique e registre eventuais erros ou conceitos subjacentes.

A *transaction log analysis*, *web log analysis* ou “análise dos percursos seguidos pelos utilizadores em cada sessão” implica a recolha ou análise de enormes quantidades de dados e uma análise demorada, tal como sucede nos testes de usabilidade com *eyetracking*, mesmo que a recolha de dados seja efetuada num curto intervalo de tempo. O risco da utilização isolada desta técnica analítica reside no facto de que os dados de utilização são, apenas, dados, pelo que é fácil fazer extrapolações, o que pode redundar em pouco rigor na análise dos dados recolhidos.

O *cognitive walkthrough* ou “**passo-a-passo cognitivo**” é uma técnica analítica que implica a intervenção de peritos em avaliação, barata e que pode ser utilizada em qualquer altura e com a frequência desejada. O avaliador delinea um conjunto de tarefas em concreto e assume o papel de um potencial utilizador interagindo com a interface do SI e anotando eventuais problemas detetados, etc. Esta técnica não permite avaliar o desempenho do SI em tempo real com utilizadores reais nem as dificuldades que um utilizador que não seja perito em avaliação possa sentir, bem

como a forma como interage com a IU, e como avalia os resultados obtidos ao recuperar informação.

A **inspeção da usabilidade** é outra técnica analítica que se baseia em participantes peritos em usabilidade, e que se aplica quer aos primeiros estágios de desenvolvimento do *design* quer a estágios de *design* mais avançados ou a SI em funcionamento. É mais económica do que os testes de usabilidade laboratoriais e é, frequentemente, orientada por tarefas do utilizador. Tal como o *cognitive walkthrough* não permite avaliar utilizadores reais, mas, sim prever os problemas que possam ocorrer aos utilizadores reais.

O *card sort* tem por objetivo perceber se um dado *website*, por exemplo, faz sentido para o utilizador, pelo que é mais útil nos casos em que o participante nunca viu o *website* ou em que aquele está a ser redesenhado. É uma técnica muito útil, não obstante, o facto de ser muito simples e rápida, bem como de necessitar de pouco investimento. É, frequentemente, usada como ferramenta de *design* na construção de árvores de menus, ou seja, é uma técnica usada no pré-lançamento de *websites*.

A **avaliação heurística** ou auditoria de usabilidade é, talvez, a técnica de avaliação de usabilidade mais utilizada⁵⁵ exatamente por ser pouco dispendiosa. É mais adequada à fase inicial de um projeto, podendo ser repetida mais adiante e é executada por um grupo de avaliadores de pequena dimensão. Tem por referência uma lista de princípios de usabilidade (*heuristics*) e permite avaliar a interface do SI ao pormenor.

A **análise de reclamações**, ou *claims analysis*, é uma técnica analítica, pelo que é realizada por um perito em usabilidade. Implica um maior nível de abstração do que a avaliação heurística ou o *cognitive walkthrough*⁵⁶. Promove uma análise em profundidade acerca da realidade e de como a mesma poderia ser diferente, os

⁵⁵ Cf. A. Blanford, S. Keith, I. Connell, H. Edwards (2004), p. 28.

⁵⁶ Idem, p.31.

avaliadores colocam hipóteses acerca das consequências do *design* e das funcionalidades disponíveis na interface para os utilizadores.

A **prototipagem de papel** é uma técnica heurística (na qual os participantes são utilizadores reais, portanto) em que os participantes executam tarefas passíveis de serem realizadas em contexto real interagindo com uma versão da interface em papel (podemos considerar um protótipo ou modelo), sendo o papel manipulado por alguém (em regra, alguém que conhece bem o *design* da interface a ser avaliada) que faz o papel do computador. Cada escolha efetuada no *écran* simulado é registada numa página em separado⁵⁷.

A **CASSM** (*Concept-Based Analysis of Surface and Structural Misfits*), ou análise baseada em conceitos de desajustes superficiais e estruturais, é uma técnica analítica que, tal como a análise de reclamações, faz uma abordagem com maior profundidade do SI a avaliar embora não permita retirar tantas ilações quanto a análise de reclamações. Centra-se na qualidade do ajuste entre os conceitos do utilizador e os conceitos do SI, analisando os conceitos com que o utilizador trabalha, os conceitos implementados no sistema e os conceitos representados na interface⁵⁸.

Para além das técnicas de avaliação da usabilidade que permitem a recolha de dados há a considerar quais os critérios que irão permitir a análise dos dados e a interpretação dos resultados do estudo de usabilidade. Têm sido vários os autores que têm enunciado critérios de avaliação da usabilidade desde 1986, sendo Nielsen (1993) e a International Standards Organization (1994) aqueles que mais vezes têm sido citados. De Shackel, em 1986, a Blanford e Buchanan, em 2003, há uma longa lista de autores que nem sempre usa a mesma terminologia: uns autores usam o termo “critério de avaliação da usabilidade” e outros usam o termo “atributo de usabilidade”. Todavia, atributos de usabilidade, mais não são do que critérios de avaliação da usabilidade, pelo que importa coligir aqueles que têm sido enunciados na literatura científica até hoje, mais concretamente de 1986 até 2003. Jeng, em

⁵⁷ Cf. C. Snyder (2003).

⁵⁸ Cf. A. Blanford, S. Keith, I. Connell, H. Edwards (2004), p. 32-33.

2005, fez um excelente inventário daquilo que designou como “atributos de usabilidade”, porém, não incluiu os cinco critérios da autoria de Blanford e Buchanan em 2003 nesse conjunto, embora os tenha mencionado. Dada a ausência de uniformidade na terminologia, recorreu-se a uma ordenação alfabética dos critérios e cronológica dos autores dentro de cada critério e criou-se a tabela 2 que consta do anexo 2 e que elucida bem a multiplicidade atrás mencionada. Como a mesma ocupa mais de duas páginas, optou-se por incluir aqui a listagem abaixo que, embora seja menos clara, evidencia, de igual modo, como são múltiplos os **critérios de avaliação da usabilidade**, os quais vão divergindo de autor para autor⁵⁹:

- ❖ Adaptabilidade - Oulanov e Pajarillo (2002)
- ❖ Afeto ou *affect*- Oulanov e Pajarillo (2002)
- ❖ Ajuda e recuperação de erros - Blanford e Buchanan (2003)
- ❖ Apreensibilidade – Shackel (1986), Booth (1989), Nielsen (1993), Hix e Hartson (1993), Kengeri, et al (1999), Blanford e Buchanan (2003)
- ❖ Aprendizagem – Furtado, et al (2003)
- ❖ Atingir objetivos - Blanford e Buchanan (2003)
- ❖ Atitude - Booth (1989)
- ❖ Atitude do utilizador - Shackel (1986)
- ❖ Capacidade de retenção ou *retainability* - Hix e Hartson (1993)
- ❖ Capacidade de ser utilizável ou *Useableness* - Gluck (1997)
- ❖ Contexto - Blanford e Buchanan (2003)
- ❖ Controlo - Oulanov e Pajarillo (2002)
- ❖ Desempenho a longo prazo - Hix e Hartson (1993)
- ❖ Eficácia – Shackel (1986), igual a *Ease of use* para Booth (1989)⁶⁰, ISO (1998), Kengeri, et al (1999)
- ❖ Eficácia da interface – Kim (2002)⁶¹
- ❖ Eficiência – Nielsen (1993), ISO (1998), Oulanov e Pajarillo (2002)
- ❖ Eficiente de usar – Brinck, et al (2002)

⁵⁹ Cf. a tabela 2 no anexo 2 para uma melhor compreensão do conteúdo desta listagem.

⁶⁰ Citado por Jeng (2005).

⁶¹ Kim, Kyunghye. 2002. A model of digital library information seeking process (DLISP model) as a frame for classifying usability problems. PhD diss., Rutgers Univ. Citada por Jeng (2005).

- ❖ Erros - Nielsen (1993)
- ❖ Experiência do utilizador - Blanford e Buchanan (2003)
- ❖ Fácil de aprender – Brinck, et al (2002)
- ❖ Fácil de recordar – Brinck, et al (2002)
- ❖ Fácil recuperação de erros - Nielsen (1993)
- ❖ Facilidade de uso – Furtado, et al (2003)
- ❖ Flexibilidade - Shackel (1986)
- ❖ Funcionalmente correto – Brinck, et al (2002)
- ❖ *Initial performance* - Hix e Hartson (1993)
- ❖ *Likeability* ou apreciabilidade - Kengeri, et al (1999)
- ❖ Memorabilidade - Nielsen (1993)
- ❖ Primeira impressão - Hix e Hartson (1993)
- ❖ Satisfação – Nielsen (1993), ISO (1998)
- ❖ Satisfação do utilizador a longo prazo - Hix e Hartson (1993)
- ❖ Subjetivamente agradável – Brinck, et al (2002)
- ❖ Taxa de erro baixa - Nielsen (1993)
- ❖ Tolerante a erros – Brinck, et al (2002)
- ❖ *User lostness* - Nielsen (1993)
- ❖ Utilidade - Booth (1989), Gluck (1997), Kengeri, et al (1999), Oulanov e Pajarillo (2002)
- ❖ Utilização de funcionalidades avançadas - Hix e Hartson (1993)

Após uma análise cuidada destes critérios, verifica-se que há conceitos que são designados por mais de um termo – sinonímia/heteronímia – e que há termos que designam conceitos diferentes ou parcialmente divergentes – polissemia – pelo que se torna difícil selecionar, objetivamente, os critérios adequados a cada caso na ausência de qualquer tipo de uniformização terminológica e de coerência concetual no conjunto da produção científica relativa a este tema.

Os critérios enunciados por Nielsen (1993), pela ISO (1998) e por Blanford e Buchanan (2003) pareceram os mais adequados a este projeto de investigação, pelo

que se irá esclarecer, apenas, com maior profundidade os conceitos desses critérios nas linhas que se seguem.

Nielsen (1993) define usabilidade como um atributo de qualidade que avalia em que medida as IU são fáceis de usar e que integra cinco (5) componentes:

- ❖ Apreensibilidade (*learnability*) – a facilidade com que o utilizador consegue concluir uma tarefa da primeira vez que toma contato com o *design*;
- ❖ Eficiência (*efficiency*) – a rapidez com que o utilizador leva a cabo as suas tarefas;
- ❖ Memorabilidade (*memorability*) – a capacidade de reestabelecer a proficiência de cada vez que um utilizador retorna ao *design* depois de ter estado um período de tempo sem contato com o mesmo;
- ❖ Erros (*errors*) – quantos erros é que o utilizador comete, qual a gravidade desses erros e com que facilidade recupera desses erros;
- ❖ Satisfação (*satisfaction*) – quão agradável é utilizar o *design*.

O termo pode, igualmente, ser usado para designar os métodos de melhoria contínua durante o processo de *design* da IU.⁶²

A norma ISO 9241-11 (1998), cuja primeira edição data de 1994, define usabilidade como “até que ponto um produto pode ser usado por um dado utilizador para atingir objetivos específicos com eficácia, eficiência e satisfação num dado contexto de uso”.⁶³ Esta norma explica, igualmente, como a usabilidade de um produto pode ser especificada e avaliada como parte de um sistema de qualidade em conformidade com a ISO 9001, bem como é que as medidas do desempenho do utilizador e da sua satisfação podem ser utilizadas para medir de que forma qualquer componente de um sistema afeta a qualidade do sistema em si.

⁶² Cf. Nielsen (1993).

⁶³ «Usability: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.»

Para Blanford e Buchanan (2003)⁶⁴ são cinco (5) os fatores de usabilidade⁶⁵ ou **critérios de avaliação da usabilidade** em bibliotecas digitais⁶⁶:

- ❖ Ajuda e recuperação de erros
- ❖ Apreensibilidade
- ❖ Atingir objetivos
- ❖ Contexto
- ❖ Experiência do utilizador

A “ajuda e recuperação de erros” permite perceber até que ponto a biblioteca digital ajuda o utilizador na utilização da mesma e a evitar e/ou a recuperar de erros cometidos durante a sua utilização. A “apreensibilidade” verifica a facilidade ou dificuldade evidenciadas pelo utilizador na aprendizagem de como usar a biblioteca digital. “Atingir objetivos” é o fator que nos permite perceber até que ponto a biblioteca digital facilita que o utilizador atinja, com eficácia e eficiência, os seus objetivos no decurso da sua utilização. O “contexto” determina até que ponto o SI se adequa ao contexto em que é usado. A “experiência do utilizador” pretende determinar o quanto o utilizador gosta de utilizar a biblioteca digital. O problema deste tipo de critérios reside no facto de os requisitos do utilizador variarem de sessão para sessão ou até dentro da mesma sessão⁶⁷. E o nível de satisfação variará em função das necessidades de informação que o utilizador quer ver satisfeitas de cada vez que interroga o SI.

Nielsen distingue entre usabilidade e funcionalidade como diferentes atributos de qualidade. E, na realidade a avaliação de bibliotecas digitais integra três facetas⁶⁸:

- ✓ Funcionalidade – opções de pesquisa, refinamento da pesquisa, limitação, capacidade de agrupar;

⁶⁴ Cf. Blanford e Buchanan (2003).

⁶⁵ Cf. Chowdhury (2011), p.155.

⁶⁶ Cf. Jeng (2005), p.49.

⁶⁷ Cf. Bollen, Luce (2002).

⁶⁸ Cf. Zhao (2010).

- ✓ Usabilidade – processo de pesquisa, navegação, apresentação do conteúdo, etiquetas;
- ✓ Acessibilidade – opções de cor, formas alternativas de apresentação do conteúdo, visualização, facilidades de navegação.

O presente projeto de investigação debruça-se sobre a usabilidade pelo que os aspetos relativos à funcionalidade e à acessibilidade não serão aqui desenvolvidos. Não obstante, listam-se, abaixo, uma série de fatores e de «funcionalidades de usabilidade para bibliotecas digitais quanto à interface do utilizador e à recuperação da informação» que Chowdhury (2004, 2011), enuncia⁶⁹ em duas das suas importantes publicações científicas:

- quanto à interface:
 - ❖ Tipo de interface (simples ou avançada)
 - ❖ Línguas da interface
 - ❖ Opções de navegação, atalhos e informação do SI
 - ❖ Funcionalidades do écran: cor, fontes tipográficas, *layout* e gráficos
 - ❖ Personalização da interface (seleção da língua, nível de pesquisa, número de registos numa página e opções de ordenação)
- quanto ao processo de pesquisa:
 - ❖ Opções para seleção
 - ❖ Meta-pesquisa
- quanto ao processamento de interrogações ou questões (*queries*)
 - ❖ Formulação e modificação das *queries* ou questões
 - ❖ Construção de conjuntos de pesquisa
 - ❖ Gravação de pesquisas
- quanto às opções de pesquisa:
 - ❖ Opções específicas de pesquisa (texto, imagem, multimédia, etc.)
 - ❖ Vários pontos de acesso ou campos de pesquisa
- quanto aos operadores de pesquisa:
 - ❖ Utilização de operadores de pesquisa

⁶⁹ Cf. Chowdhury (2011) – tradução livre e adaptação de conteúdo das p. 156-157.

- ❖ Interrogações em linguagem natural
- quanto à manipulação de resultados:
 - ❖ Formato(s) de visualização dos registos
 - ❖ Número de registos visualizados
 - ❖ Navegação na lista de registos
 - ❖ Marcação dos registos
 - ❖ Opções de ordenação
 - ❖ Impressão, exportação e envio por *email* de registos
- quanto à ajuda ao utilizador:
 - ❖ Adequação e utilidade
 - ❖ Ajuda contextualizada
 - ❖ Consistência da terminologia, do *design* e do *layout*
 - ❖ Correção linguística

Todas estas funcionalidades podem ter impacto no desempenho do utilizador no decurso de um processo de pesquisa, pelo que acabam por ter implicações na própria usabilidade do SI. No entanto, dado que funcionalidade e usabilidade são conceitos distintos, a lista acima enunciada não poderá ser considerada para uma avaliação da usabilidade no sentido estrito do termo. Pese embora o facto de a usabilidade poder ser aparente ou inerente⁷⁰, isto é, a usabilidade aparente está mais ligada à visualização da IU e a usabilidade inerente está ligada ao aspeto dinâmico e funcional da IU.

Neste projeto de investigação que se centra na perspetiva do utilizador, **serão utilizados os seguintes critérios de avaliação da usabilidade:**

- ✚ Ajuda e recuperação de erros
- ✚ Atingir objetivos
- ✚ Eficiência
- ✚ Satisfação

⁷⁰ Cf. Kurosu, Kashimura (1995).

 *User lostness*

A maioria dos estudos de usabilidade de bibliotecas digitais ocorreu já no século XXI, havendo a reportar dois estudos levados a cabo no século passado antes do famoso artigo de Saracevic (2000)⁷¹ que ainda foi publicado no último ano do séc. XX na *Library Trends*.

Após uma análise da produção científica disponível, ressaltam os nove estudos de usabilidade em bibliotecas digitais que se seguem por ordem cronológica:

- ❖ 1998 - DeLiver – Bishop (1998), Neumann, Bishop (1998)
- ❖ 2000 - ACM Digital Library, NCSTRL (Networked Computer Science Technical Reference Library), NZDL (New Zealand Digital Library) – Theng, Mohd-Nasir, Thimbleby (2000)
- ❖ 2000 - PDL (Perseus Digital Library) – Marchionini (2000);
- ❖ 2001, 2003 - DLESE (Digital Library for Earth System Education) – Sumner, Dawe (2001); Sumner, Khoo, Recker, Marlino (2003)
- ❖ 2002 – DLF (Digital Library Federation) – Covey (2002)
- ❖ 2002 – Hospital de Londres de grande dimensão – Adams, Blanford (2002)
- ❖ 2003 - NSDL (National Science Digital Library) – Sumner, Khoo, Recker, Marlino (2003)
- ❖ 2004 - NCSTRL (Networked Computer Science Technical Reference Library) – Hartson, Shivakumar, Pérez-Quñones (2004)
- ❖ 2005 - eLibraryHub – Theng, Chan, Khoo, Buddharaju (2005)
- ❖ 2010 – EUROPEANA⁷² – Dobрева, Chowdhury (2010)⁷³.

O estudo sobre a **DeLiver**⁷⁴ envolveu mil e novecentos (1900) estudantes de cursos de pós-graduação de quatrocentas e vinte (420) faculdades e utilizou uma

⁷¹ Cf. Saracevic (2000).

⁷² Acessível na Internet: «<http://www.europeana.eu/>».

⁷³ DOBREVA, M. AND CHOWDHURY, S., 2010. A user-centric evaluation of the Europeana digital library. In Springer-Verlag, pp. 148-157. Apresentado em 23/06/2010 na ICADL'10.

⁷⁴ Cf. Neumann, Bishop (1998), Bishop (1998).

combinação de vários métodos: *transaction log analysis*, inquérito (questionário, entrevista e grupos de foco) e teste de usabilidade. Incidiu sobre a acessibilidade e levou os autores a concluir que os fatores de acessibilidade têm grande influência sobre o uso, que a recolha uniforme de dados e o seu cruzamento é essencial, e que é importante o contínuo desenvolvimento de métricas adequadas e este tipo de SI digitais que fazem parte da infraestrutura informacional em mudança.

O estudo sobre a **ACM Digital Library**, a **NCSTRL** e a **NZDL**⁷⁵, levado a cabo em 2000⁷⁶, envolveu quarenta e cinco (45) estudantes de cursos de 1º ciclo (sendo distribuídos quinze por cada biblioteca digital) e incidiu no *design* e na estrutura, tendo como métodos o questionário e a avaliação heurística e como critério a satisfação. Os autores concluíram que este estudo deverá ser alargado a outras áreas científicas (estudantes e pessoal que não sejam da computação) para se ter uma compreensão mais abrangente do propósito das bibliotecas digitais. O estudo permitiu identificar uma elevada percentagem de *User lostness* nas três (3) bibliotecas digitais – cerca de 73% no conjunto das três -, sendo detetados diferentes níveis de *lostness* ou sensação de se estar perdido no SI digital. Os intervenientes no estudo referiram que, na sua maioria, as bibliotecas digitais não proporcionavam suficiente informação que lhes permitisse compreender a estrutura das mesmas – 67% na NCSTRL, 53% na NZDL e 47% na ACMDL.

A **Perseus Digital Library (PDL)**⁷⁷ foi desenvolvida no âmbito de um projeto cujo planeamento começou em 1985, e que se encontra em contínuo desenvolvimento desde 1987, que teve por objetivo criar uma biblioteca digital para as humanidades que desse acesso ao maior conjunto de fontes de informação para essa área. Gary Marchionini⁷⁸ efetuou um estudo de avaliação da PDL, em 2000, em que usou uma abordagem multifacetada que incluía múltiplos métodos e técnicas de investigação,

⁷⁵ A NZDL é a biblioteca digital que lançou o *software open source* que mantém o seu nome original e que se notabilizou como um dos principais suportes deste tipo de SI digitais, o “Greenstone Digital Library”.

⁷⁶ Cf. Theng, Mohd-Nasir, Thimbleby (2000).

⁷⁷ Cf. <http://www.perseus.tufts.edu/>

⁷⁸ Cf. Marchionini (2000).

tais como o *think-aloud*, a *transaction log analysis*, as entrevistas, a análise documental e a análise de aprendizagem.

O estudo sobre a **DLESE**⁷⁹ e a **NSDL**⁸⁰ envolveu trinta e sete (37) professores de vários níveis de ensino e dois (2) bibliotecários distribuídos por cinco (5) grupos de foco, e incidiu sobre o *design*, com o objetivo de identificar as expectativas e os requisitos para a utilização de coleções de bibliotecas digitais em ambiente de sala de aula, concluindo que os participantes sentem que as suas perceções são influenciadas por fatores como a qualidade do conteúdo, o *design*, a publicidade. Muito importante foi a constatação de que os utilizadores com perfil de ensino sentem necessidade de bibliotecas digitais mais rápidas, com conteúdos de melhor qualidade, mais fáceis de usar, com dados qualitativos que partam da população de utilizadores e que promovam competências de raciocínio crítico.

Os **estudos** de usabilidade⁸¹ conduzidos por **membros da DLF**⁸² foram caracterizados, num documento publicado em Janeiro de 2002, por Denise Troll Covey no estudo abrangente⁸³ que elaborou para a Digital Library Federation como tendo o(s) seguinte(s) perfil/perfis:

- Avaliação heurística;
- Estudo de *card sorting*;
- Questionário;
- Prototipagem;
- Protocolos *think-aloud* conduzidos com prototipagem.

O estudo conduzido num grande **Hospital de Londres**⁸⁴ por Adams e Blanford, em 2002, teve como participantes setenta e três (73) profissionais de saúde distribuídos

⁷⁹ Cf. Sumner, et al. (2003) e Sumner, Dawe (2001).

⁸⁰ Cf. Sumner, et al. (2003).

⁸¹ Ter em atenção que se reportam, na generalidade, a bibliotecas híbridas.

⁸² Cf. Covey (2002).

⁸³ Cf. as p.51-52 desta tese para mais pormenores.

⁸⁴ Cf. Jeng (2005), p. 51.

por grupos de foco aos quais foram realizadas entrevistas e incidiu sobre a acessibilidade. De entre os participantes, 50% eram enfermeiros e os outros 50% eram médicos, gestores, consultores e membros do departamento de TI. O conceito de biblioteca digital não foi definido⁸⁵, surgindo referências a bases de dados especializadas na área biomédica, como a Medline, como sendo bibliotecas digitais que foram objeto do estudo, pelo que o interesse das conclusões deste estudo é relativo para esta tese. Conclui que as bibliotecas digitais aumentam o acesso à informação, mas que a sua usabilidade é fraca e que implicam dispêndio de tempo que não existe no dia-a-dia dos clínicos. Foi notado um hiato de competências em TI entre os profissionais mais novos e os profissionais de faixas etárias mais avançadas, o que permite concluir – partindo da descrição efetuada pelos próprios autores⁸⁶ - que muitos dos problemas de “usabilidade” relatados para as “bibliotecas digitais” em análise fossem, na realidade, problemas de literacia digital.

O estudo realizado para a **NSDL** coincidiu com o estudo da **DLESE**⁸⁷ e foi feito em conjunto com aquele. A National Science Digital Library é um programa de investigação alargado financiado pela National Science Foundation (NSF) dos EUA que tem como objetivo construir uma biblioteca digital para a educação em ciência, matemática, engenharia e tecnologia⁸⁸.

O estudo elaborado para a **NCSTRL**⁸⁹ incidiu sobre o *design*, a interface e a funcionalidade, tendo recorrido à técnica de investigação designada como inspeção da usabilidade que foi realizada por três (3) peritos em usabilidade. Concluiu que o *design* estava orientado para a funcionalidade e que a terminologia utilizada na NCSTRL incluía gíria profissional estando centrada no *designer* e não no utilizador.

⁸⁵ Cf. Adams, Blanford (2002), p. 60.

⁸⁶ Idem, p. 65.

⁸⁷ Cf. o parágrafo sobre o estudo da DLESE nas p. 67-68 desta tese.

⁸⁸ Cf. Arms, et al. (2002).

⁸⁹ Cf. Hartson, et al (2004).

O estudo levado a cabo para a **eLibraryHub**⁹⁰ centrou-se na estrutura de navegação e na organização da informação tendo como técnicas o questionário e a análise de reclamações e como critérios a eficácia, a utilidade e a satisfação. O questionário foi respondido por cento e sete (107) participantes (de entre estudantes e pessoal) e a análise de reclamações envolveu seis (6) participantes. Dado que este estudo aplicou um misto de métodos quantitativos e qualitativos, foi crucial o cruzamento de dados. Os resultados apontaram para a necessidade de efetuar várias melhorias no SI. É considerado aconselhável que o trabalho futuro do primeiro estudo (questionário) envolva uma amostra mais alargada de participantes quer em número quer em diversificação de perfis, e que o do segundo estudo (análise de reclamações) poderia ter vantagem na diversificação dos participantes do ponto de vista demográfico, apesar dos bons resultados obtidos com o mesmo. É, igualmente, apontada a hipótese de introduzir outras técnicas de investigação como a entrevista.

O estudo conduzido sobre a **EUROPEANA**⁹¹, por Dobрева e Chowdhury⁹², foi divulgado na conferência ICADL'10 que decorreu de 21 a 25 de Junho de 2010 na Gold Coast na Austrália. Usou uma combinação de grupos de foco com tecnologia *eyetracking* para levar a cabo estudos de usabilidade na Biblioteca Digital Europeia. Incidiu sobre vários aspetos como a facilidade de uso e o carácter intuitivo do SI (sobretudo para novos utilizadores), identificação de necessidades futuras do utilizador e definição de estilos de uso de um protótipo (jovens utilizadores), expectativas do utilizador incluindo a confiança no SI, variações entre utilizadores de vários países. O estudo envolveu participantes selecionados oriundos do Reino Unido, da Holanda, da Itália e da Bulgária. As impressões dos utilizadores foram divididas em três categorias e a discussão dos resultados fez-se com base nessas categorias: primeiras impressões (expectativas), impressões profundas (experiência

⁹⁰ Cf. Theng, et al. (2005).

⁹¹ O European Digital Library Project foi lançado pela Comissão Europeia (CE) em 2005, e é financiado pela CE⁹¹, tendo dado origem à biblioteca digital europeia EUROPEANA que disponibiliza recursos digitais na *World Wide Web* em parceria com as agências bibliográficas nacionais de todos os países da União Europeia, entre outras instituições nacionais e internacionais, agregando no seu seio, em 2013, outros quarenta e sete (47) projetos. A EUROPEANA constitui, em 2013, uma infraestrutura digital que agrega a biblioteca digital, o museu digital e o arquivo digital europeus.

⁹² Cf. Dobрева, Chowdhury (2010).

do utilizador), e impressões marcantes (se o utilizador voltará a usar o SI e quais as alterações que são necessárias). As autoras do estudo realçaram o benefício do cruzamento dos dados de *eyetracking* com dados recolhidos por outros métodos e/ou técnicas de investigação para o estudo do comportamento do utilizador⁹³.

Dois (2) destes estudos deram origem a modelos ou técnicas de estudo da usabilidade em bibliotecas digitais – o da **PDL** e o da **DLESE/NSDL**. Para além daqueles, existem mais três (3) modelos e técnicas de estudo da usabilidade em bibliotecas digitais⁹⁴:

- ❖ CIAO! aplicado à KMODDL;
- ❖ NeLCD – National Electronic Library for Communicable Disease;
- ❖ MEDLIS – um modelo de usabilidade em bibliotecas digitais.

O **CIAO! (Context, Interaction, Attitude and Outcomes)**⁹⁵ é um modelo de avaliação que foi desenhado por Jones, et al. para avaliar a usabilidade de software educacional e que foi utilizado por Pan, et al. para avaliar a KMODDL (Kinetics Models for Design Digital Library), em 2004⁹⁶. O CIAO! incide no contexto, na interação, na atitude e nos resultados e utiliza questionários, entrevistas, *software logs* e *web logs*, análise documental, observações em sala de aula e registo em vídeo. Pan, et al. pretenderam avaliar a integração dos recursos da KMODDL (área da matemática) no ensino secundário e no ensino superior de 1º ciclo e determinar até que ponto eles facilitam ou melhoram a aprendizagem. O estudo concluiu que havia melhorias a efetuar em termos de usabilidade na biblioteca digital e que a usabilidade influencia diretamente a aprendizagem, ou seja, quanto mais usável é um SI digital que se utiliza no processo de ensino-aprendizagem melhores resultados são alcançados na aprendizagem. Uma aprendizagem eficaz induz uma experiência positiva subjetivamente. Mas, a boa usabilidade não é indispensável a uma boa aprendizagem e uma experiência positiva nem sempre indica boa usabilidade, dado

⁹³ Cf. Chowdhury (2011).

⁹⁴ Idem.

⁹⁵ Cf. Jones, et al. (1999).

⁹⁶ Cf. Pan, et al. (2004).

que alguns utilizadores desta faixa etária se entretiveram com aspetos da simulação que mais não eram do que *bugs* do SI e, assim, dispersaram a sua atenção do processo de ensino-aprendizagem.

O modelo de avaliação da **NeLCD** (National Electronic Library for Communicable Disease)⁹⁷ tem por objetivo determinar a eficácia das bibliotecas digitais da área biomédica e dos *websites* de informação em saúde na mudança de atitude e de comportamento dos médicos na relação médico-doente e na tomada de decisão clínica e respetiva comunicação, bem como dos restantes profissionais de saúde e do público. O modelo combina *transaction log analysis* com questionários (prévio e posterior) para avaliar o impacto da NeLCD no conhecimento, na atitude e no comportamento do utilizador. Os estudos de avaliação da usabilidade levados a cabo nesta biblioteca digital com este modelo de avaliação concluíram que há evidência de que as bibliotecas digitais médicas podem mudar o conhecimento e as atitudes dos utilizadores, e que estes estudos são essenciais para assegurarem que as bibliotecas digitais médicas satisfazem as necessidades dos utilizadores e que contribuem para a melhoria da qualidade dos cuidados de saúde.

O **MEDLIS** é o primeiro⁹⁸ modelo genérico destinado à avaliação da usabilidade das bibliotecas digitais, e foi desenvolvido e testado por Chowdhury, McMenemy e Poulter em 2008⁹⁹. Os autores defendem que um estudo de usabilidade deve começar por colocar questões de âmbito genérico, devendo ter-se em conta quem poderá responder da melhor forma a essas questões – o que envolve utilizadores e pessoal da área da informação. Não obstante, há informação que só poderá ser obtida junto dos utilizadores, tal como: o motivo pelo qual usam o serviço, facilidade de uso do mesmo e o nível de satisfação do utilizador com o mesmo. O modelo apresentado com a designação MEDLIS tem início com duas questões: qual o tipo de serviço a avaliar, e quem é que o fornece? As respostas podem ser dadas por membros da equipa de profissionais que asseguram o *design* e a gestão diária do SI e o

⁹⁷ Cf. Madle, et al. (2003).

⁹⁸ Conclusão retirada pelos autores após a revisão da literatura efetuada.

⁹⁹ Cf. Chowdhury, McMenemy, Poulter (2008).

desenvolvimento e o financiamento do mesmo. Neste caso, poderia ser aplicado o inquérito (questionários, entrevistas e grupos de foco) e a análise documental para estudar aspetos relativos à gestão como a origem, os objetivos, relatórios de mercado, gestão estratégica, políticas governamentais, e o modelo económico. Ainda para obter resposta às primeiras duas questões, através do mesmo perfil de pessoal ligado à gestão do serviço, aplicar-se-ia a avaliação heurística para os aspetos relacionados com o *design* e o conteúdo. O modelo estrutura-se em quatro (4) colunas com as seguintes designações:

- ✓ Questões genéricas;
- ✓ Quem pode responder;
- ✓ O que estudar;
- ✓ Como.

A última coluna “Como” indica os métodos a serem utilizados em cada caso. Dado que o modelo é bastante complexo, ficando bem mais claro se explicado visualmente num diagrama¹⁰⁰, não se opta pela descrição integral do mesmo e faz-se uma listagem completa dos métodos/técnicas de recolha de dados enunciados no mesmo:

- ✓ O inquérito (questionários, entrevistas / grupos de foco);
- ✓ A análise documental;
- ✓ A avaliação heurística;
- ✓ A análise de mercado;
- ✓ Observação;
- ✓ *Log analysis*;
- ✓ Método experimental.

A generalidade destes estudos de usabilidade centra-se no utilizador. E é, também, essa a preocupação deste projeto de investigação – avaliar os arquivos e as bibliotecas digitais sob a perspetiva do utilizador.

¹⁰⁰ Cf. Chowdhury (2011), p.164, figura 8.1

«When I use a word', Humpty Dumpty said in a rather scornful tone,

'it means just what I choose it to mean – neither more or less.'»

Lewis Carroll (1832–1898),

Through the Looking Glass (1871), Chapter 6

Capítulo 2

Os sistemas de informação digitais e o acesso à informação: conceitos e contextualização científica

Como o próprio título indica, este capítulo da tese destina-se a contextualizar cientificamente o seu conteúdo e a clarificar, sucintamente, os conceitos que nela são usados e a aceção em que são referidos.

2.1. Os sistemas de informação e os repositórios digitais

Relembrando o ponto um da introdução¹⁰¹ desta tese, os sistemas de informação em suporte digital surgem, apenas, após a 2ª Grande Guerra Mundial a meio do século

¹⁰¹ Cf. p. 21-22.

passado. Grace Hopper proferiu, em 5 de Fevereiro de 1987, na Ohio State University em Columbus nos EUA, uma frase que viria a ficar para a história das ciências da computação: «Life was simple before World War II. After that we had systems»¹⁰². Na realidade, já tínhamos sistemas e já tínhamos SI, mas não os tínhamos em suporte eletrónico ou digital. E isso teve um enorme impacto. O aparecimento da informática e a introdução dos computadores na nossa vida organizacional e pessoal quotidiana teve um impacto incomensurável só comparável ao aparecimento da escrita, mas com um desenvolvimento muitíssimo mais acelerado e uma progressão exponencial. E são os SI digitais que são alvo deste projeto, mais concretamente os arquivos digitais e as bibliotecas digitais. E aqui entra em jogo mais um termo: os repositórios digitais. São, portanto, estes os conceitos que importa definir neste ponto 2.1.

O **conceito de SI** tem sido alvo de várias definições por parte de vários autores, as quais têm vindo a variar ao longo do tempo e à medida que as características dos SI evoluem. Beynon-Davies, em 2009, define SI¹⁰³ como um sistema de comunicação entre pessoas que implica o armazenamento, o processamento, a distribuição e a utilização da informação, e em que o resultado principal é a informação em si, a qual suporta sistemas de atividade nas organizações¹⁰⁴. Pode alargar-se um pouco mais o conceito e dizer-se que a informação suporta a atividade humana e até a própria existência.

Os SI digitais incluem uma variada tipologia como diretórios, *websites*, motores de pesquisa, portais temáticos, bibliotecas digitais, arquivos digitais, museus digitais... O termo “repositório” (sejam eles institucionais, temáticos, científicos, de dados, etc.) é o termo que mais se lê e ouve na área da ciência da informação e ciências

¹⁰² <http://www.cs.yale.edu/homes/tap/Files/hopper-wit.html>

¹⁰³ Cf. Beynon-Davies (2009), p.2 - «An information system is a system of communication between people. Information systems are systems involved in gathering, processing, distributing and using information. The key output of an information system is of course information, which is used to support activity systems in organisations.»

Cf. Beynon-Davies (2013), p.15.

¹⁰⁴ É certo que Beynon-Davies refere esta definição no contexto de um livro sobre SI de negócios. Todavia, a definição é dada logo no início como sendo aplicável a “sistemas de informação” de modo genérico.

afins, desde fins do séc. XX, a seguir a bibliotecas digitais. Já o mesmo não se pode dizer, sobretudo cientificamente, do termo “arquivos digitais”, o que parece um contrassenso¹⁰⁵. A terminologia continua, ainda, um pouco ambígua na literatura científica, apesar de o conceito de “biblioteca digital” se ter estabilizado, consensualmente, neste termo já há alguns anos depois de uma deambulação do termo “biblioteca virtual” para “biblioteca eletrónica” e desse para este. Os arquivos digitais têm sido mais vezes tratados por repositórios do que as bibliotecas digitais, já para não falar nos museus digitais. Não querendo entrar numa discussão que seria muito interessante, mas que não teria cabimento aqui, dado o objetivo do capítulo e da tese em si, considera-se que “**repositório digital**” é um termo mais específico de “sistema de informação digital” e que é um termo mais genérico de “arquivo digital” e de “biblioteca digital”. Assim, o termo “repositório digital” integra-se, como subcategoria, no termo “sistema de informação digital” e subdivide-se em vários tipos de repositórios, entre os quais se encontram os arquivos digitais e as bibliotecas digitais. Se considerarmos as várias “facetas” possíveis, começaremos a perceber que poderão ser institucionais ou de outro tipo sob esse ponto de vista, de dados científicos ou outro, etc. O **arquivo digital** pode ser considerado como um repositório que reúne conteúdos não publicados como a designada “literatura cinzenta”, ou versões prévias de documentos em fase de publicação como *drafts* (rascunhos) ou *pre-prints*, conteúdos relativos a dados científicos que podem ser reutilizados, resumindo, conteúdos criados/armazenados/organizados/divulgados ou comunicados/preservados/avaliados/eventualmente eliminados/reeditados e reutilizados no decurso de uma atividade organizacional ou pessoal. Esta definição evoca o ciclo de vida da informação, no caso, da informação digital, dado que a reedição (não propriamente a reutilização...) é uma realidade da era digital¹⁰⁶. O gráfico apresentado na página seguinte pretende clarificar as fases da vida da informação digital. Onde se indica “partilha” poderia constar “disseminação” ou “divulgação” ou “comunicação”. Mas, partilha parece ser um termo mais adequado à realidade dos repositórios digitais sejam eles abertos ou fechados.

¹⁰⁵ Mais estranho, ainda, após o lançamento da Open Archives Initiative (OAI), em 1999, e após a Convenção de Santa Fé em 22 de Maio do mesmo ano.

¹⁰⁶ Ainda que os manuscritos fossem anotados e rasurados, por vezes, a reedição de documentos em arquivo só se tornou uma prática comum e sequencial na era digital.



Figura 5 – Ciclo de vida da informação digital

O conceito de **biblioteca digital** tem sido alvo das mais variadas definições nas últimas décadas, as quais oscilam consoante o contexto (científico e não só) de quem as propõe e que refletem diferentes visões. Não obstante, tal como sucedeu na terminologia com a escolha do termo que representa o conceito, começa a chegar-se a algum consenso em torno das características de uma biblioteca digital.

Em 1994, a Digital Libraries Initiative (DLI) financiada¹⁰⁷ pela NSF¹⁰⁸, NASA e DARPA, definia a biblioteca digital do seguinte modo:

«Digital libraries basically store materials in electronic format and manipulate large collections of those materials effectively. Research into digital libraries

¹⁰⁷ FOX, Edward A., 2005. Digital Libraries Initiative (DLI) Projects 1994-1999. In: *Bulletin of the American Society for Information Science and Technology* [online]. 31 January 2005. Vol. 26, no. 1, pp. 7-11. [Accessed 11 December 2013]. DOI 10.1002/bult.135. Available from: <http://doi.wiley.com/10.1002/bult.135>.

¹⁰⁸ National Science Foundation dos EUA.

is research into network information systems, concentrating on how to develop the necessary infrastructure to effectively mass-manipulate the information on the Net. The key technical issues are how to search and display desired selections from and across large collections.»

Em 1997, Michael Lesk define as bibliotecas digitais, muito sucintamente, como coleções organizadas de informação digital¹⁰⁹. Em 2012, o mesmo autor refere no resumo de um dos seus artigos:

«Progress in digital libraries is described in four areas: technical, economic, legal and social issues. Originally technological obstacles were dominant, but they have generally been overcome by progress in computers, networks, and algorithms. Economic issues have also faded, although *open access* questions still bedevil us. Surprisingly, libraries are now faced with more serious legal obstacles than first expected; no one today can start a digital library effort without thinking about copyright. Social issues are likely to be the next set of challenges.»¹¹⁰

De facto, os aspetos sociais são demasiado importantes para a compreensão das bibliotecas digitais, de tal forma que são vistas como sistemas sociotécnicos altamente complexos.

Em 1999, a Digital Library Federation (DLF) definia, desta forma, uma biblioteca digital:

«Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.»¹¹¹

¹⁰⁹ Cf. Lesk (1997), p.1.

¹¹⁰ Cf. Lesk (2012), p.592.

¹¹¹ <http://old.diglib.org/about/dldefinition.htm>

Esta definição da DLF inclui como característica da biblioteca digital a prestação de serviços, o que a distingue de um arquivo digital ou de uma coleção digital. Mas, será essa uma exclusividade da biblioteca digital? Não poderão o arquivo digital e o museu digital prestar serviços na *World Wide Web*?! Parece ser evidente que sim. Mas, esta distinção da biblioteca digital pelos serviços que oferece está relacionada com a visão de quem propõe a definição. No mesmo ano, Christine Borgman¹¹² observava que, enquanto os investigadores viam as bibliotecas digitais como conteúdo digital recolhido e reunido para usufruto da comunidade, os bibliotecários viam as bibliotecas digitais como instituições ou serviços. Na realidade, ambas as visões estão corretas e plasmam duas potenciais facetas de uma mesma realidade. Em 2002, Borgman, Solvberg e Kovács¹¹³, olham para as bibliotecas digitais como novas formas de organizações/instituições da área da informação, como sistemas de recuperação da informação multimédia ou como sistemas de informação que suportam a criação, o uso e a pesquisa de conteúdo digital. Concordemos que essa definição se pode aplicar aos arquivos digitais, e mais genericamente, aos repositórios digitais. Em 2000, William Arms apresentara a seguinte definição de biblioteca digital:

«An informal definition of a digital library is a managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network.»¹¹⁴

Em 2004, no texto de descrição da rede DELOS¹¹⁵, é referido o seguinte:

«Digital Libraries represent a new infrastructure and environment that has been made possible by the integration and use of a number of IC technologies, the availability of digital content on a global scale and a strong

¹¹² Christine L. Borgman. "What are digital libraries? Competing visions" *Information Processing & Management* 35.3 (1999): 227-243.

¹¹³ http://wwwold.sztaki.hu/conferences/deval/presentations/DELOSWorkshop4OnEval_report.pdf

¹¹⁴ Cf. Arms (2000), p.2.

¹¹⁵ http://www.delos.info/index.php?option=com_content&task=view&id=299&Itemid=26

demand from users who are now on-line. They are destined to become an essential part of the information infrastructure in the 21st century.»

Nos últimos anos, agora que a definição do conceito começa a estabilizar, surgiram autores das ciências da computação a defenderem a adoção de um novo termo para representar o conceito de forma a separá-lo da noção de biblioteca e de arquivo alegando que as bibliotecas e os arquivos digitais não surgiram como continuidade daqueles serviços de informação na era digital, que não era essa a intenção de quem os criou... E assim surge o termo “ciber-infraestrutura”. Esse termo parece, apenas, muito mais abrangente do que os outros dois e do que o termo repositórios digitais. É um termo que representa um conceito muito mais vasto do que os repositórios digitais e, se quem “criou” os arquivos e as bibliotecas digitais não pensou em bibliotecas e em arquivos, quem teve a visão que inspirou os criadores dos repositórios digitais – Vannevar Bush¹¹⁶ – pensou nas bibliotecas e nos arquivos pessoais¹¹⁷. Quem sabe organizar, representar e recuperar informação melhor do que os profissionais de informação?! O caminho da sociedade do conhecimento e da sociedade em rede não pode ser feito de forma sectária se quisermos ser bem-sucedidos. Teremos de o percorrer de forma cooperante, adquirindo novas competências, partilhando conhecimentos e trabalhando em equipa para o progresso da sociedade em que vivemos, fazendo um esforço mútuo para compreendermos as visões e os modelos de cada área e para nos adaptarmos à era digital pugnando pela qualidade dos sistemas de informação.

Concluindo, a biblioteca digital pode ser entendida como um repositório que reúne conteúdos publicados recebidos/armazenados/organizados/divulgados ou partilhados/preservados. Qualquer um destes tipos de repositórios digitais pode ter várias outras características adicionais decorrentes do seu contexto – serem abertos ou fechados, públicos ou privados, gerais ou especializados, etc.

¹¹⁶ Cf. p. 21 desta tese.

¹¹⁷ Cf. Bush (1945).

2.2. A interface do utilizador e a pesquisa por assunto

A interface do utilizador (IU) para a recuperação da informação relaciona-se com várias áreas científicas, sobretudo, com as ciências da computação, com a ciência da informação e com a psicologia. A recuperação da informação constitui parte essencial do processo de acesso à informação, pelo que o *design* da IU é um elemento central na comunicação entre os seres humanos e as máquinas, ou seja, entre utilizadores/pesquisadores e sistemas de recuperação da informação (SRI). Dado que o processo de pesquisa e recuperação da informação é, em si mesmo, um processo dinâmico e interativo, conseguir desenhar uma IU eficaz no estabelecimento de uma comunicação adequada entre os seres humanos e as máquinas é uma tarefa bastante difícil. A interação homem-máquina (*human-computer interaction* em língua inglesa, cuja sigla reconhecida é HCI) tornou-se num campo de investigação e desenvolvimento vital nas ciências da computação, uma vez que quer o acesso à Internet quer a informação nela disponível aumentam, exponencialmente, a cada dia que passa. Hoje em dia, quem efetua pesquisas é cada vez mais o utilizador final e não um mediador ou intermediário profissional. As pesquisas dos utilizadores finais são executadas, preferencialmente, em texto integral, e não tanto em bases de dados exclusivamente bibliográficas, rankings de relevância são disponibilizados pelos SRI na visualização dos resultados e a navegação é facilitada por ligações entre as fontes de informação. Todos estes fatores reforçaram, desde o início da década de 80 do séc. XX, a necessidade de avaliar a usabilidade dos SRI, primeiro com testes de usabilidade e mais tarde com novas técnicas de avaliação da usabilidade¹¹⁸.

As IU são canais de comunicação entre seres humanos e sistemas, logo, as **IU para recuperação da informação** estabelecem comunicação entre pesquisadores humanos e SRI. A este propósito, pode citar-se Marti Hearst (1999):

«Information seeking is an imprecise process. When users approach an information access system they often have only a fuzzy understanding of how

¹¹⁸ Cf. o ponto 1.2. desta tese.

they can achieve their goals. Thus the user interface should aid in the understanding and expression of information needs. It should also help users formulate their queries, select among available information sources, understand search results, and keep track of the progress of their search.»¹¹⁹

E podem, igualmente, ser citados Marchionini e Komlodi (1998):

«Understanding the information-seeking process and developing systems and strategies for supporting it are central goals of information science. Research in the organization and communication of information is best informed by studies of the interactions among people and external information sources. However, information technology has advanced so rapidly in the second half of the twentieth century that it dominates research and development in information seeking, and the linchpin of this interactivity is the user interface. The interactions among human physical, cognitive, and affective subsystems and the external world are defined by the juxtaposed boundaries where these physical and conceptual constructs meet. Such conjunctions of boundaries are called interfaces. Interfaces serve as the communication channels through which information seeking proceeds. Defining and building interfaces that support information seeking is thus a fundamental problem in information science and there is a rich history of work that may be found in the human factors and human computer interaction (HCI) literatures as well as in the information science literature.»¹²⁰

Mas, a subjetividade é uma característica humana – as pessoas têm diferentes motivações e diferentes comportamentos consoante a sua idade, género, personalidade, aptidões e competências, educação, contexto social e cultural. Tendo em conta estes fatores, desenhar IU universais e eficazes parece ser um objetivo

¹¹⁹ Marti Hearst - *User interfaces and visualization*, p. 257.

¹²⁰ MARCHIONINI, G. AND KOMLODI, A., 1997. Design of Interfaces for Information Seeking. *Annual Review of Information Science and Technology (ARIST)*, 33, pp. 89-130.

difícil de atingir. Uma IU eficaz não deve ser perceptível enquanto é usada, deve facilitar o nosso trabalho sem nos fazer perder muito tempo a aprender a lidar com ela e a ajustá-la. Deve disponibilizar um *feedback* informativo, permitir uma fácil reversão de ações, suportar um ponto de controlo interno, reduzir o carregamento da memória interna e fornecer IU alternativas para utilizadores iniciais e para utilizadores avançados¹²¹. Escolher uma IU adequada deveria ser uma ação intuitiva facilmente levada a cabo por cada tipo de utilizador. Portanto, o *design* de *software* requer flexibilidade. Decidir qual a informação a ser visualizada na IU constitui uma importante decisão a ser tomada pelos *designers* de IU. Som, apresentações 3D, animação e vídeo são utilizados pelos *designers* para facilitar a interação homem-máquina. As funcionalidades da IU incluem¹²²:

1. Cor, gráficos, imagens
2. Animação, ampliação e redução (*zooming*), vídeo
3. Tamanho de visualização variável (baixa a alta resolução)
4. Janelas redimensionáveis pelo utilizador
5. Barras de ferramentas
6. Sons, música, entrada e saída de voz
7. Rato, teclas de cursor, écran tátil, *stylus*, novos dispositivos de entrada
8. Métodos alternativos de entrada
9. Línguas internacionais
10. Usabilidade universal

Manipulação direta, telepresença e realidade virtual são algumas das técnicas utilizadas pelos *designers* de IU capazes de mudarem a nossa vida quotidiana.

As IU têm três conhecidos atributos de qualidade: a usabilidade, a universalidade e a utilidade, pelo que a avaliação da usabilidade tem uma parte importante no processo de *design* da IU¹²³.

¹²¹ Idem, p. 258-259.

¹²² Ben Shneiderman, Catherine Plaisant - *Designing the user interface*, p. 207.

¹²³ Melody Ivory, Marti Hearst - The state of the art in automating usability evaluation of user interfaces, 2001, p. 470.

Concluindo, a interface do utilizador para recuperação da informação pode ser definida como um canal de comunicação através do qual se desenrola o processo de pesquisa e recuperação da informação.

2.3. Os sistemas de organização do conhecimento

A organização da informação, e posteriormente do conhecimento, surge por motivos do foro administrativo, mas também culturais e educacionais. Os primeiros manuscritos de gestão administrativa guardados/armazenados acabam por constituir arquivos. Mas, surgem, também, obras filosóficas, historiográficas, literárias que preservadas em conjunto acabam por formar bibliotecas. Os primeiros documentos que registaram mensagens escritas ou textos, e que chegaram até nós, encontravam-se nos mais variados suportes:

- tabuinhas de argila (Suméria);
- tabuinhas de madeira ou de marfim enceradas;
- baixos-relevos em mármore, granito, etc.;
- papiro;
- pergaminho.

A biblioteca de Ebla na Suméria, cerca de 4500 a.C., é, talvez, a biblioteca mais antiga que se conhece. Nela foram descobertos dezassete mil (17000) fragmentos de tabuinhas, isto é, cerca de quatro mil (4000) documentos. Estes documentos estavam repartidos por duas salas de um palácio, sendo os da mais pequena todos de carácter económico e os que se encontravam na maior (18 m²) textos administrativos, legais, históricos, religiosos e linguísticos. A forma das tabuinhas estava relacionada com o tipo de informação contida, o que permitia a sua organização em estantes de madeira. Na lombada ou borda de cada uma encontrava-se uma legenda que identificava o seu conteúdo. São inúmeros os textos literários, e encontraram-se trinta e dois (32) dicionários sumério-eblaítas. Há, ainda, silabários e gramáticas. Na cidade residiam cerca de dezoito mil (18000) escribas. É evidente a presença, nos vestígios desta

biblioteca, das funções técnicas mais essenciais da área profissional da informação: classificação, cotação, ordenação e conservação. O arquivo de Lagash na Suméria – 2000 a.C. -, tal como a biblioteca de Ebla, evidencia cuidados com o armazenamento dos documentos, permitindo a sua preservação – protegiam as tabuinhas de argila da falta de humidade, dado que, na sua maioria, aquelas não se encontravam cozidas e o clima era bastante seco, inserindo-as em bancos, cestas e jarras. A forma mais comum, na época, de destruição das cidades pelo inimigo – os incêndios - é que, ironicamente, acabou por aumentar a sua durabilidade... Na escavação arqueológica do arquivo de Mari na Suméria – 1700 a.C. - foram encontradas tabuinhas de madeira encerada e, também, rolos de pele. Os recipientes (cestos) onde se conservavam os documentos eram identificados com etiquetas de argila seguras por fios de vime onde constavam a classe dos documentos contidos e a sua data (dia, mês e ano). Também, foram encontradas jarras a servirem de recipientes de armazenamento. Na escavação arqueológica da biblioteca de Asurbanipal – 612 a.C. -, entre as tabuinhas de argila aí descobertas em trinta mil (30000) fragmentos, encontrava-se um conjunto onde constava o “Poema de Gilgamesh”, o poema escrito mais antigo da humanidade. As tabuinhas copiadas no escritório do palácio deste rei assírio, incluíam colofão. Nalguns constava, no meio de um texto algo vaidoso escrito em seu próprio nome, a menção: “os símbolos da escrita, todos os que foram inventados, escrevi-os em tabuinhas, ordenei-as em séries¹²⁴, coleccionei-as e coloquei-as no meu palácio para minha real contemplação e leitura”. Estes colofões incluíam, também, maldições (o que não sucede nos babilónicos), tal como mais tarde se vem a verificar nos manuscritos medievais.

O colofão foi outra inovação mesopotâmica e outro importante contributo desta civilização para o livro e para as bibliotecas. É na Mesopotâmia que se encontram as raízes da arquivística e da biblioteconomia. As bases da preservação e conservação dos documentos, da sua organização e da arrumação para uma recuperação posterior, são lançadas nesta civilização. Os locais onde se encontram os documentos são de pequenas dimensões, o que permitia uma organização dos documentos. Havia água dentro desses locais para manter o nível de humidade do ar (de forma a evitar rachas

¹²⁴ Curiosa expressão para a altura: “ordenei-as em séries”.

na argila) e, também, para apagar aquelas tabuinhas cujo conteúdo já não interessava permitindo que fossem reescritas¹²⁵. Também havia fornos para cozer as tabuinhas cujo conteúdo se pretendia preservar. Havia vários tipos de estantes e de recipientes para armazenamento e organização dos documentos, mantendo juntos os que se reportavam a um mesmo assunto. Os recipientes eram etiquetados, o que permitia a sua rápida localização. Os documentos eram identificados na sua lombada ou borda por inscrições convencionadas, e os documentos em várias peças também eram devidamente correlacionados, incluíam colofões que descreviam as obras e identificavam, por vezes, o escriba e o proprietário¹²⁶. Surge, também, pela primeira vez a portada ou rosto nas tabuinhas enceradas, as quais faziam lembrar, pela sua forma, os códices ou os livros. Os primeiros catálogos ou listas de obras, também surgiram na Mesopotâmia - descreviam, possivelmente, o conteúdo de estantes ou de salas da biblioteca. Pode considerar-se que são estes os primórdios da organização do conhecimento.

Com o aparecimento da imprensa, no século XV, surge a bibliografia¹²⁷, embora já existam “relações de códices” guardados em bibliotecas particulares de casas nobres e de mosteiros na Idade Média em Portugal. Às bibliografias especializadas, as mais comuns no dito período humanista da história da bibliografia (sécs. XVI-XVIII)¹²⁸, vem juntar-se, no século XVI, a primeira bibliografia geral. Em 1547, em plena Inquisição, por ordem da Real Mesa Censória é editado o primeiro volume dos índices expurgatórios que são, pode considerar-se, as primeiras bibliografias portuguesas. Porém estes índices não tinham por objetivo preservar a memória escrita e sim listar obras proibidas para apreender os bens dos seus proprietários, assassinar aqueles e destruir as obras censuradas, ou seja, o resultado último era a eliminação da memória escrita que desagradava ao Tribunal do Santo Ofício. Daí que

¹²⁵ Afinal a reedição dos textos que a informática nos trouxe com os programas de processamento de texto e que nos fez repensar o ciclo de vida da informação na era digital, já era uma prática em arquivos da Antiguidade pré-clássica, ainda que de forma diversa.

¹²⁶ O rigor de identificar autor e proprietário (que, em princípio, seria o produtor).

¹²⁷ Em 1494 surge a primeira obra que pode ser considerada como uma bibliografia - “Liber de scriptoribus ecclesiasticis” por Johann Tritheim - e que refere cerca de 7000 obras. Pela data, percebe-se que se trata de uma relação de Incunábulos.

¹²⁸ Cf. Malclés (1977).

se considere que a bibliografia portuguesa criada com o propósito de preservar para memória futura uma relação das obras publicadas existentes numa determinada data só tenha surgido em 1741 com o primeiro tomo da obra intitulada “Bibliotheca lusitana historica, critica e cronologica”¹²⁹ da autoria de Diogo Barbosa de Machado que é considerado o pai da bibliografia portuguesa.¹³⁰ É da preocupação com a elaboração de bibliografias para inventariar as obras publicadas e o conhecimento, e do conseqüente estabelecimento de regras e técnicas decorrentes dessa prática, que acabam por surgir os primeiros sistemas de organização do conhecimento na segunda metade do séc. XIX – as classificações. A taxonomia de Linneu, de 1735, é anterior às primeiras classificações e decorreu da preocupação naturalista, equivalente à bibliológica, de inventariar os seres vivos existentes na Terra. As bases mínimas de técnicas de catalogação e de classificação surgem, pois, em fins do séc. XVIII. Segundo Louise-Noëlle Malclés, de 1790 a 1810 decorre o período bibliofílico e após 1810¹³¹ entra-se no período técnico e profissional começando a época técnica em 1914, o que acaba por só ocorrer após 1920 quando o Instituto Internacional de Bibliografia (IIB) retomou a sua atividade após a Primeira Grande Guerra¹³². Com a construção das primeiras classificações, na segunda metade do séc. XIX, a descrição dos documentos deixa de ser, apenas, relativa à sua forma e passa a debruçar-se, igualmente, sobre o seu conteúdo - a análise e representação da informação, a indexação propriamente dita, tem início nesta altura.

Os **sistemas de organização do conhecimento (SOC)** têm tido uma presença algo discreta nos repositórios digitais, apesar de serem uma peça essencial na representação e na recuperação da informação. Considera-se, nesta tese, que aquele termo representa, de forma muito abrangente, todas as linguagens ou vocabulários

¹²⁹ https://bdigital.sib.uc.pt/bduc/Biblioteca_Digital_UCFL/digicult/UCFL-CFE-9-1_4/globalItems.html

¹³⁰ Conrad Gessner é considerado por muitos “o pai da bibliografia” a nível internacional por ter publicado em 1545 a “Bibliotheca Universalis”, uma bibliografia que inclui referências a cerca de 15000 obras manuscritas e impressas de cerca de 3000 autores.

¹³¹ 1810 marca a publicação do “Manuel du libraire et de l’amateur des livres” por Jacques-Charles Brunet.

¹³² Cf. op. cit.

controlados e não-controlados, incluindo novas realidades¹³³ como redes semânticas e ontologias, utilizados na representação da informação. E será nessa aceção que será invocado ao longo da mesma¹³⁴.

F. Wilfrid Lancaster explicou, muito sinteticamente, no diagrama reproduzido¹³⁵ na figura 6, a importância dos SOC no processo de acesso à informação / recuperação da informação. Se a estratégia de pesquisa incluir, na interrogação ao SRI, os mesmos descritores que foram usados para representar a informação aquando da indexação, a probabilidade de recuperar resultados relevantes para a necessidade de informação do utilizador é muito maior. Mas, para que tal aconteça é necessário que os SOC utilizados na representação da informação estejam disponíveis e acessíveis no SI digital para consulta do utilizador aquando da submissão da *query* ao SRI. As bases de dados bibliográficas e de texto integral de acesso fechado (algumas têm disponibilizado acesso aberto com limitação de acesso a alguns aspetos como os SOC) têm investido na presença dos SOC e algumas são quase imbatíveis na qualidade da recuperação da informação como é o caso da Medline, não obstante as críticas dos adeptos da indexação automática¹³⁶. A criação, gestão e manutenção de SOC de qualidade é dispendiosa, e até morosa, mas compensa na fase de recuperação da informação, uma vez que orienta o utilizador na pesquisa e que lhe poupa imenso tempo na execução da mesma. Porém, continua a haver vozes discordantes que alegam ser suficiente a utilização de um excelente motor de pesquisa como o Google

¹³³ Com origem nas ciências da computação, e sem a colaboração da ciência da informação, como é o caso das ontologias.

¹³⁴ Nesta tese, os SOC constituem o termo preferencial para linguagens de indexação e para linguagens documentais. Nem todos os autores estarão de acordo com esta opção, mas este não será o local mais oportuno para encetar essa discussão, sob pena de ocorrer um desvio do foco do projeto de investigação aqui descrito.

¹³⁵ Cf. Lancaster (2003).

¹³⁶ Gerard Salton, em 1972, não conseguiu melhor desempenho com o SMART do que os testes efetuados por Lancaster em 1969 ao MEDLARS. Em 1992, Hersh testa o Saphire e no resumo de um seu artigo pode ler-se: «An analysis of SAPHIRE, an experimental information retrieval system featuring automated indexing and natural language retrieval, was performed on MEDLINE references using data previously generated for a MEDLINE evaluation. Compared with searches performed by novice and expert physicians using MEDLINE, SAPHIRE achieved comparable recall and precision. While its combined recall and precision performance did not equal the level of librarians, SAPHIRE did achieve a significantly higher level of absolute recall. SAPHIRE has other potential advantages over existing MEDLINE systems. Its natural language interface does not require knowledge of MeSH, and it provides relevance ranking of retrieved references.»

para se aceder a tudo o que se encontra num repositório digital aberto. E isso pode até funcionar muito bem se o acesso é feito via Google. Mas, e quando se acede diretamente ao repositório? Neste projeto de investigação procuram-se respostas para questões relacionadas com estas preocupações.

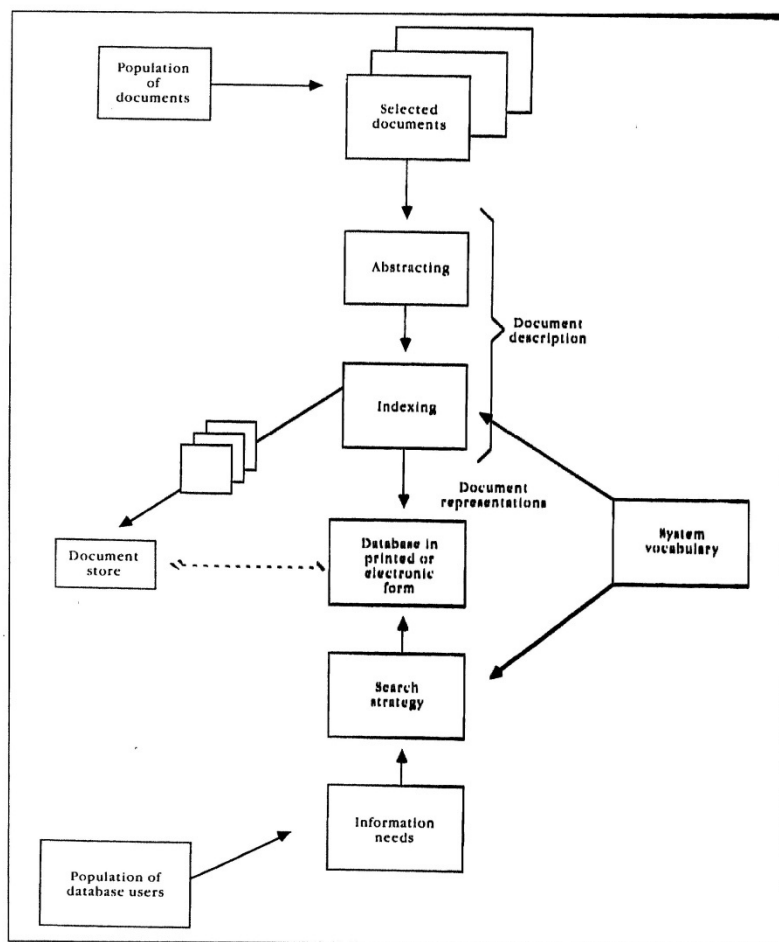


FIGURE 1

The role of indexing and abstracting in the larger information retrieval picture.

Figura 6 – O papel central dos SOC no processo de recuperação da informação visto por Lancaster – imagem reproduzida a partir da p. 2 do seu livro *Indexing and abstracting in theory and practice*.

Gail Hodge¹³⁷ realizou, em 2000, um relatório para a DLF em que identificava os SOC¹³⁸ presentes nas bibliotecas digitais de então, tendo enunciado uma listagem

¹³⁷ Cf. Hodge (2000). Acessível na Internet: <http://www.clir.org/pubs/reports/pub91>

¹³⁸ Hodge definiu os SOC da seguinte forma: «The term *knowledge organization systems* is intended to encompass all types of schemes for organizing information and promoting knowledge management¹.

daquilo que entendia serem sistemas de organização do conhecimento, a qual foi retomada por Hjørland, em 2007, a propósito de um capítulo acerca da importância da semântica na ciência da informação e, mais concretamente, na organização do conhecimento¹³⁹:

- ✚ Cabeçalhos por assunto
- ✚ Classificações
- ✚ Dicionários
- ✚ Ficheiros de autoridade
- ✚ Gazetas
- ✚ Glossários
- ✚ Ontologias
- ✚ Planos de classificação
- ✚ Redes semânticas
- ✚ Taxonomias
- ✚ *Thésauri*

Em 2005, foi publicado um importante documento para o ensino da ciência da informação (CI), sob o título “European Curriculum Reflections on Library and Information Science Education”¹⁴⁰, onde ficaram definidas todas as áreas disciplinares que integram esta ciência por cerca de 150 cientistas (da CI) provenientes de toda a Europa. No capítulo 7, dedicado à organização do conhecimento é enunciada uma listagem de tipos de SOC que segue abaixo:

« In the narrow meaning (within LIS) the kind of KOS, which are considered include:

Knowledge organization systems include classification schemes that organize materials at a general level (such as books on a shelf), subject headings that provide more detailed access, and authority files that control variant versions of key information (such as geographic names and personal names). They also include less-traditional schemes, such as semantic networks and ontologies. Because knowledge organization systems are mechanisms for organizing information, they are at the heart of every library, museum, and archive.»

¹³⁹ Hjørland, B. (2007). Semantics and knowledge organization, ARIST, 41, 367-405. Acessível na Internet: https://www.academia.edu/1912560/Semantics_and_knowledge_organization

¹⁴⁰ <http://www.asis.org/Bulletin/Dec-06/EuropeanLIS.html>

- Enumerative classification systems a la Dewey, UDC and LCC
- Facet analytic systems a la Bliss 2nd ed.
- Subject headings like LCSH
- Systems based on free text searches
- Thesaurus based systems
- Bibliometric maps
- Algorithms in search engines
- Archival systems (based on the principle of provenance)
- Ontologies
- Semantic networks
- "Topic maps"
- etc.

These systems may all be regarded as kinds of semantic tools providing selection of concepts and information about their semantic relations. »¹⁴¹

Estes SOC têm origens semânticas nuns casos e sintáticas noutras, e mistas nalguns outros. As linguagens controladas integram linguagens combinatórias¹⁴² que expressam as relações semânticas entre os termos, ainda que possam ter apresentação sistemática, e as linguagens categoriais¹⁴³ que expressam as relações sintáticas entre os termos, ainda que (usualmente) incluam um índice alfabético (ou uma caixa de diálogo pesquisável por termo na versão digital), e que podem ser hierárquicas ou facetadas. Ao longo do tempo foram surgindo tipos mistos de linguagens – a Classificação Decimal Universal (CDU)¹⁴⁴ é o primeiro exemplo de incorporação de características combinatórias aquando da inclusão de tabelas auxiliares e de sinais/símbolos de combinação de notações já na segunda metade do séc. XX. Mais

¹⁴¹ Cf. Lørring, Kajberg, eds. (2005), p.142.

¹⁴² Este tipo de linguagem controlada surge na segunda metade do séc. XX em consequência do aparecimento da informática e da teoria do *Uniterm* de Mortimer Taube (1951), tendo o princípio combinatório origens mais remotas, nos anos 30 do mesmo século, na primeira classificação facetada – a “Colon Classification” de Ranganathan.







¹⁴³ Este tipo de linguagem controlada surge na segunda metade do séc. XIX, com precursores em Sir Francis Bacon (1561-1626), Hegel (1770-1831) e Brunet (1780-1867).

¹⁴⁴ A CDU surge, em 1893, como uma adaptação da Dewey Decimal Classification (DDC), que tinha sido publicada em 1876, da autoria de Henri La Fontaine e Paul Otlet. A adaptação da DDC foi autorizada por Dewey a pedido de Otlet e La Fontaine.

tarde, surgem os classaurus como um misto de linguagem categorial e combinatória. O Architecture & Art Thésaurus¹⁴⁵ do The Getty Research Institute incorporou facetas num thésaurus criando um vocabulário extremamente útil para quem indexa e para quem pesquisa na área da Arte. Mais recentemente, surgem novos SOC como as ontologias e as redes semânticas, realidades de uma era da informação e do conhecimento suportada por SI digitais que alicerçam a atividade humana a nível organizacional e a nível pessoal.

Um vocabulário controlado é, basicamente, uma lista de autoridade, tal como bem refere Lancaster,¹⁴⁶ que incorpora uma estrutura semântica onde são expressas relações de equivalência, hierárquicas e associativas - no caso das linguagens combinatórias -, ou uma estrutura sintática/sistemática - no caso das linguagens categoriais ou classificatórias.

No âmbito do trabalho laboratorial deste projeto de investigação (e por motivos que se prendem com as características dos repositórios digitais selecionados para a amostra bem como com a frequência com que cada tipo de SOC surge nas fontes de informação secundárias e terciárias) foram considerados os seguintes SOC nas aceções expostas nas definições que são apresentadas nos parágrafos seguintes:

-  Classificação
-  Índice de assuntos
-  Nuvem de assuntos
-  Ontologia
-  Taxonomia
-  Tesouro¹⁴⁷

Por **classificação** entende-se uma linguagem ou vocabulário controlado que expressa relações sintáticas entre os termos¹⁴⁸, de forma hierárquica ou por facetas,

¹⁴⁵ <http://www.getty.edu/research/tools/vocabularies/aat/index.html>

¹⁴⁶ Op. cit., p. 19.

¹⁴⁷ Termo em português para *thésaurus*.

representando por notações (compostas por combinações alfanuméricas e de símbolos) os conceitos identificados, aquando da primeira fase da indexação, na análise concetual. Os descritores ou termos preferenciais são todas as notações.

Por **índice de assuntos** entende-se uma fonte de informação secundária composta por uma lista ordenada de termos ideográficos em que a cada termo corresponde uma referência que aponta para a sua localização.

Por **nuvem de assuntos**¹⁴⁹ entende-se uma forma de visualização de uma lista ou agregação de etiquetas (*tags*) ou palavras-chave que pode ser ordenada e que coloca em destaque aquelas que têm maior número de ocorrências, geralmente, através do tamanho da letra. Quanto maior é a letra, maior número de ocorrências tem a palavra-chave apresentada na etiqueta. Este SOC surgiu no contexto da Web 2.0 e está relacionado com a indexação social ou *social tagging*¹⁵⁰. Mais abaixo, na figura 7, reproduz-se um exemplo.

Por **ontologia** entende-se um SOC que estrutura as fontes de informação na *World Wide Web* e que expressa as relações semânticas entre os termos e as propriedades dos termos, permitindo que linguagens de anotação descrevam, automaticamente, as fontes de informação presentes na *World Wide Web*. Tem imensas parecenças com o tesauro, distinguindo-se do mesmo pela explicitação das propriedades dos termos, frequentemente, em estruturas lógico-sintáticas.

Por **taxonomia** entende-se uma linguagem ou vocabulário controlado que expressa relações semânticas hierárquicas entre termos que correspondem a categorias apresentadas numa estrutura sistemática. Cada classe ou categoria da árvore sistemática constitui um descritor¹⁵¹.

¹⁴⁸ Podem ser vistas como relações semânticas hierárquicas entre estruturas sintáticas...

¹⁴⁹ *Tag cloud* ou *word cloud* em inglês.

¹⁵⁰ Também conhecida como *folksonomy*.

¹⁵¹ Este tipo de SOC foi o primeiro a surgir sendo anterior às classificações. Surge no meio científico, nas ciências naturais, com a taxonomia de Linneu, em 1735 ao ser publicada a sua obra "Systema Naturae". Cf. <http://upload.wikimedia.org/wikipedia/commons/2/29/SN-p837.jpg>

2.4. Os sistemas de recuperação da informação

A análise e representação da informação e a recuperação da informação são duas facetas de um mesmo processo, como ficou bem claro no diagrama de Lancaster reproduzido na figura 6, e constituem o núcleo central da ciência da informação. Os sistemas de recuperação da informação (SRI) são uma peça essencial no acesso à informação e o seu desempenho é determinante para o sucesso de todo o SI. O investimento no tratamento da informação só faz sentido se a ela se puder aceder de forma eficaz e, desejavelmente, célere. O fim último de qualquer profissional de informação é o de satisfazer as necessidades de informação do utilizador ou cliente, consoante se encontre numa organização pública ou privada. O propósito de qualquer SI deve, em última análise, ser esse mesmo.

2.4.1. A pesquisa e recuperação da informação como área científica

A pesquisa e recuperação da informação¹⁵³ é uma área científica que partilha terreno quer com a ciência da informação (CI) quer com as ciências da computação (CC), embora com abordagens diferentes. Na CI, os SRI são estudados pelo lado da pesquisa da informação e do comportamento do utilizador no acesso à informação e, conseqüentemente, aos SRI em si mesmos. Na CC, os SRI são vistos pelo lado do SI, da construção do SRI e do processo de recuperação da informação por dentro do próprio SRI. Em conjunto com a análise e representação da informação, constitui o núcleo central da CI.

A pesquisa e recuperação da informação é uma área da ciência da informação que se subdivide em três (3) áreas com ligações interdisciplinares a outras ciências:

- ✚ a pesquisa da informação (estudada, também, nas ciências da vida)¹⁵⁴;
- ✚ a recuperação da informação (partilhada com as ciências da computação);

¹⁵³ *Information Seeking & Retrieval (IS&R)*.

¹⁵⁴ E com interesse para a psicologia, a sociologia e até as ciências políticas.

✚ e o comportamento informacional (com ligações às ciências do comportamento).

Qualquer destas três áreas criou, ao longo dos anos, uma comunidade científica específica e sólida com as suas próprias conferências e publicações científicas, e grupos de trabalho em associações científicas. Todavia, a sua atividade tem-se desenvolvido em nicho sem que haja uma constante colaboração, como seria desejável para um maior progresso na área de *IS&R*, o que se percebe, sobretudo, entre a pesquisa da informação (PI) e a recuperação da informação (RI). Entre a pesquisa da informação e o comportamento informacional a colaboração é mais evidente, o que se deverá, decerto, ao facto de ambas as áreas serem das ciências sociais e humanas, tendo, portanto, um *modus operandi* similar. A agenda separada da PI e da RI deve-se à sua diferente contextualização científica e nota-se, sobretudo, na maior visibilidade da produção científica da RI, o que terá a ver com o ritmo mais acelerado da produção científica em CC e um maior número de conferências dedicadas a RI. Como resultado, surgiram diferentes modelos e terminologias algo divergentes em PI e em RI. Recentemente, o aparecimento de conferências comuns como a “Information Interaction in Context” (IiC) tem tentado ultrapassar esta limitação da área de *IS&R* com a construção de uma comunidade científica mais abrangente que pode ter inúmeras vantagens, por exemplo, na partilha de projetos de investigação.

2.4.2. A pesquisa da informação

A pesquisa da informação tem raízes nas “ciências” documentais (a biblioteconomia e a arquivística como práticas profissionais com origens milenares) e desbravou caminho numa era analógica em que era bem mais complexo e trabalhoso encontrar informação relevante para as necessidades de informação dos utilizadores. Das bibliografias e dos catálogos de livreiros aos catálogos em fichas, dos cartões perfurados aos microfilmes e às disquetes 5^{1/4} e 3^{1/2} (suportes magnéticos), existe toda uma história de SRI manuais e semiautomáticos até se chegar aos SRI

automáticos com os OPAC, as bases de dados e bancos de dados *online* via Dialog, BRS, Orbit e outras empresas fornecedoras de serviços na área da informação, os CD-ROM (e outros suportes óticos), e, finalmente, com a *World Wide Web*. Os SRI automáticos acabaram por extravasar a Internet e invadir as Intranets organizacionais sendo imbuídos nos SI organizacionais como os SAP e outros ERP. A informação foi descrita e pesquisada por profissionais de informação e investigadores, sobretudo, ao longo de séculos e hoje em dia essa tarefa alargou-se a toda a população numa verdadeira democratização da informação e do conhecimento originada pela sociedade global das últimas décadas. As ferramentas da Web 2.0, a *web* social, permitem que o cidadão comum crie, edite, partilhe e publique informação com maior celeridade e com uma repercussão exponencial comparativamente com as da Web 1.0. Os blogues, as redes sociais, as *folksonomies* ou indexação social, são realidades da Web 2.0 que permitem a descrição da informação por qualquer um e a sua difusão rápida por toda a rede global. A Web 1.0 permitia a edição, publicação e comunicação/difusão mas não ao mesmo ritmo e com muito menor impacto. E a indexação social surge, apenas, na *web* social. Perante este cenário, a pesquisa da informação que é vista pelas perspetivas social e cultural, para além da perspetiva do SI, tem um enorme e aliciante desafio pela frente quanto a investigação e desenvolvimento.

Järvelin e Ingwersen (2004) definem **pesquisa da informação** como sendo percecionada como um processo evolutivo:

«Information seeking has been understood as a process in which the actor's understanding of his or her tasks or problems, information needs, relevance criteria, and the available information space evolve. »¹⁵⁵

Pode considerar-se este conceito como sendo mais abrangente do que o de recuperação da informação, dado que o termo “recuperação” indicia que se encontra algo que foi armazenado e é, habitualmente, aplicado a suportes digitais. Na pesquisa

¹⁵⁵ Cf. Järvelin, Ingwersen (2004).

ou busca da informação¹⁵⁶ procura-se informação em qualquer tipo de suporte, seja analógico ou digital, formal ou informalmente, e esse processo não é estudado apenas dentro da CI¹⁵⁷, segundo Hjørland¹⁵⁸, sendo estudado em todos os seres vivos por cientistas da biologia¹⁵⁹. Marchionini, em 1995, caracteriza a pesquisa da informação e distingue-a da recuperação da informação de uma forma bastante clara:

« Information seeking is a process in which humans purposefully engage to change their state of knowledge. The term search will be used to mean the behavioral manifestation of humans engaged in information seeking and will also be used to describe the actions taken by computers to match and display information objects. The term "information seeking" is preferred to "information retrieval" because it is more human-oriented and open-ended. Retrieval implies that the object must have been "known" at some point; most often, whoever "knew" it organized it for later "knowing" by themselves or someone else. Seeking connotes the process of acquiring knowledge; it is more problem-oriented; the solution may or may not be found. (...) Retrieval is applicable to database management and most applied problems, but seeking is closer to question-answering or learning. Information seeking is a fundamental human process closely related to learning and problem solving. (...) The ability to locate and apply information is an important component of what it means to be literate. (...) Information seeking, like learning, is a fundamental and high level cognitive process. Information seeking is often part of learning or problem solving, but it is also distinct. »¹⁶⁰

No âmbito deste projeto de investigação procedeu-se a uma avaliação da presença de SOC na interface do utilizador de SI digitais do tipo arquivo digital e biblioteca digital aquando do processo de recuperação da informação. Ou seja, o foco do projeto de investigação, e desta tese, na área científica da pesquisa e recuperação da

¹⁵⁶ Mais adiante, esclarece-se a distinção entre os termos "searching" e "seeking" por Wilson (1999).

¹⁵⁷ *Library and Information Science* (LIS) em língua inglesa.

¹⁵⁸ http://www.iva.dk/bh/core%20concepts%20in%20lis/articles%20a-z/information_searching.htm

¹⁵⁹ Cf. Bell (1991).

¹⁶⁰ Cf. Marchionini (1995), p. 5-6.

informação, encontra-se nos SRI, pelo que a atenção se centrará mais na recuperação da informação.

2.4.3. A recuperação da informação

A recuperação da informação (RI) tem origens¹⁶¹ que se confundem com a génese da CI¹⁶², sendo esta posterior e tendo raízes naquela, e com a história da CC¹⁶³ como ciências e que remontam ao período pós 2ª Grande Guerra Mundial e à posterior aplicação à sociedade em geral dos avanços tecnológicos conseguidos com o financiamento que foi concentrado na investigação científica para fins militares no decurso daquela. O termo “recuperação da informação” foi criado por Calvin Moers, em 1951, que lhe atribuiu a seguinte definição:

«Information retrieval is the name for the process or method whereby a prospective user of information is able to convert his need for information into an actual list of citations to documents in storage containing information useful to him. It is the finding or discovery process with respect to stored information. It is another, more general, name for the production of a demand bibliography. Information retrieval embraces the intellectual aspects of the description of information and its specification for search, and also whatever systems, techniques, or machines that are employed to carry out the operation. Information retrieval is crucial to documentation and organization of knowledge. »¹⁶⁴

A RI estabeleceu-se como área científica dinâmica com congressos, publicações científicas, inúmeros cientistas e projetos de investigação na comunidade científica internacional. Insere-se na CI e na CC com uma maior percentagem de cientistas na CC, dado que se reportou sempre à informação digital e aos SI digitais, ou seja, à

¹⁶¹ Cf. p.21-22, 83 desta tese.

¹⁶² A CI data de inícios dos anos 60 do século XX (1961-1962). Cf. <http://comminfo.rutgers.edu/~tefko/SaracevicInformationScienceELIS2009.pdf>

¹⁶³ <http://www.wolframalpha.com/docs/timeline/computable-knowledge-history-5.html>

¹⁶⁴ Cf. Moers (1951), p.25. Cf. Saracevic (1996), p.44.

recuperação da informação automática. Os SRI manuais e semiautomáticos são, historicamente, objeto de estudo da pesquisa da informação no âmbito da CI. Da recuperação da informação em citações e em descrições bibliográficas¹⁶⁵ passou-se à recuperação da informação em texto integral e com o auxílio de sistemas inteligentes como o aplicado pela Amazon há muitos anos atrás e pela Google, o que veio alargar o universo de utilizadores e tornar o papel dos intermediários no processo de RI menos imprescindível. Na última década e meia, com o movimento de acesso livre ao conhecimento, a informação armazenada para ser recuperada, nomeadamente na Internet, passou a incluir o texto integral da dita “literatura cinzenta” e dos *pre-prints* e *post-prints* de publicações científicas de todo o mundo, pelo que se passou a dispor de repositórios digitais agregadores do Conhecimento e acessíveis durante as 24 horas do dia e os 365 dias do ano. Ou seja, passámos do documento para a informação e da informação para o conhecimento em pouco mais de meio século. E estamos a cumprir a visão de Vannevar Bush. Isto deu uma dimensão completamente diferente aos SRI automáticos e uma importância sem paralelo à recuperação da informação como área científica e disciplinar, dado que a mesma se dedica aos SRI automáticos e digitais. Neste sentido vão as afirmações de Van Rijsbergen, já em 1979, no excerto que se segue:

«Information retrieval is a wide, often loosely-defined term, but in these pages I shall be concerned only with automatic information retrieval systems. Automatic as opposed to manual and information as opposed to data or fact. Unfortunately the word information can be very misleading. In the context of information retrieval (IR), information, in the technical meaning given in Shannon’s theory of communication, is not readily measured. In fact in many cases, one can adequately describe the kind of retrieval by simply substituting “document” for “information”. Nevertheless, “information retrieval” has become accepted as a description of the kind of work published by Cleverdon, Salton, Sparck Jones, Lancaster and others. A perfectly straightforward definition along this line is given by Lancaster: “Information

¹⁶⁵ Apelidada de “recuperação de documentos” por alguns.

retrieval is the term conventionally, though somewhat inaccurately, applied to the type of activity discussed in this volume. An information retrieval system does not inform (i.e., change the knowledge of) the user on the subject of his inquiry. It merely informs on the existence (or non-existence) and whereabouts of documents relating to his request”. This specifically excludes Question-Answering systems as typified by Winograd and those described by Minsky. It also excludes data retrieval systems such as used by, say, the stock exchange for on-line quotations. »¹⁶⁶

Manning, Raghavan e Schütze, definem recuperação da informação de uma forma bastante mais abrangente que reflete a realidade da sociedade da informação e do conhecimento do final da primeira década do séc. XXI:

«Information retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers). »

Destas definições depreende-se que o termo **recuperação da informação** significa extração de informação a partir do conteúdo de coleções (digitais). E para que se concretize a extração da informação é necessário um SRI, uma base de dados ou um banco de dados ou um repositório digital (que, em última análise, pode ser a própria Internet) onde estejam armazenados registos cujo conteúdo tenha sido indexado e um ficheiro inverso pesquisável. Ora para que tudo isso fosse exequível foram surgindo, ao longo das décadas, modelos de recuperação da informação que começaram por ser de navegação no início da década de 90 do século passado, até chegarmos aos modelos de recuperação estruturados. Ricardo Baeza-Yates sintetizou num diagrama os vários modelos de recuperação da informação como se pode ver na figura 8 abaixo reproduzida. Cada um deles trouxe funcionalidades de pesquisa aos SRI como a navegação plana, os operadores booleanos, os rankings de relevância, etc.

¹⁶⁶ Cf. Van Rijsbergen (1979), p.1.

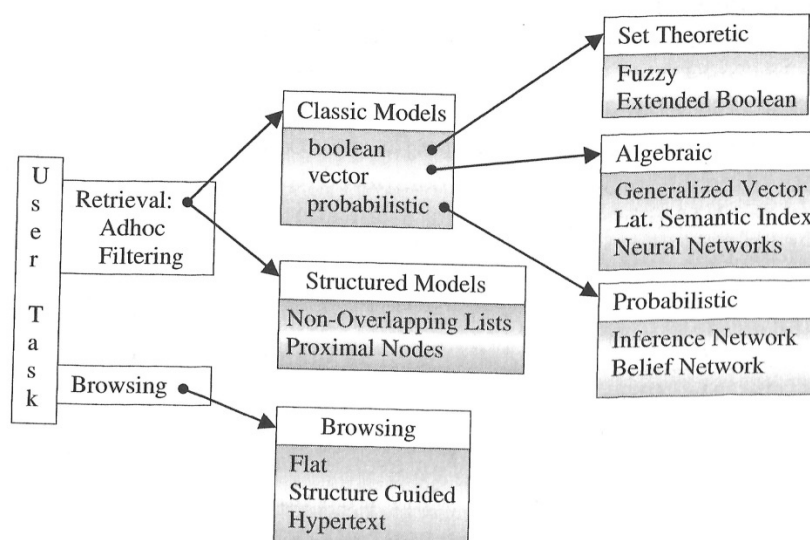


Figure 2.1 A taxonomy of information retrieval models.

Figura 8 – Taxonomia de modelos de recuperação da informação por Ricardo Baeza-Yates e Berthier Ribeiro-Neto¹⁶⁷

No âmbito desta tese, interessam as **funcionalidades de pesquisa dos SRI dos repositórios digitais**, as quais não costumam ser profusas, mas que podem incluir uma ou várias das dezassete (17) funcionalidades¹⁶⁸ a seguir enunciadas:

- Ajuda e documentação de apoio
- Extensão da pesquisa (dados bibliográficos, texto integral, diferentes tipos de documentos ou de ficheiros)
- Índices
- Interfaces do utilizador simples e avançadas
- Operadores booleanos
- Pesquisa *case sensitive* (letras maiúsculas e/ou minúsculas, acentuação, etc.)
- Pesquisa lateral (*related items*)
- Pesquisa multimédia
- Pesquisa por campo
- Pesquisa por data ou período cronológico (campo numérico)

¹⁶⁷ Reproduzido a partir de Baeza-Yates, Ribeiro-Neto (1999), p. 21.

¹⁶⁸ Cf. Smith (2000).

- Pesquisa por extração de frase e operadores de proximidade
- Ranking de relevância
- Refinamento da pesquisa inicial
- Técnicas de visualização
- Tradução da língua
- Truncatura
- Vocabulário controlado (*thésaurus*, etc.)

As funcionalidades que foram consideradas no projeto de investigação descrito nesta tese foram os “**índices**” (mais concretamente, os índices por assunto) e o “**vocabulário controlado**”, na designação de Smith, que ficaram definidos no ponto 2.3. relativo aos SOC¹⁶⁹.

2.4.4. O comportamento informacional e os estudos do utilizador

O comportamento informacional centra-se no utilizador, dando ênfase às necessidades de informação, ao uso da informação e ao processo cognitivo, pelo que se relaciona com as ciências do comportamento e com a sociologia. Os primeiros estudos do utilizador são anteriores a 1970 e, em regra, estão orientados para o SI e não para o utilizador propriamente dito. Os estudos do utilizador, na verdadeira aceção da palavra, começam a ser publicados na década de 70 do século XX, sendo a década de 80 aquela em que começam a surgir importantes definições na área do comportamento informacional como a separação do termo composto “*information needs and use*” que agregava dois conceitos num só termo: necessidade de informação e uso da informação. O comportamento informacional tem sido alvo de várias teorizações que resultaram em modelos teóricos da autoria de vários investigadores, de entre os quais se destaca o contributo de T. D. Wilson. Segundo Wilson, deve distinguir-se entre comportamento informacional, comportamento de busca de informação, comportamento de pesquisa de informação e comportamento de uso da informação:

¹⁶⁹ Cf. p. 95-97.

«**Information Behavior** is the totality of human behavior in relation to sources and channels of information, including both active and passive information seeking, and information use. Thus, it includes face-to-face communication with others, as well as the passive reception of information as in, for example, watching TV advertisements, without any intention to act on the information given.

Information Seeking Behavior is the purposive seeking for information as a consequence of a need to satisfy some goal. In the course of seeking, the individual may interact with manual information systems (such as a newspaper or a library), or with computer-based systems (such as the World Wide Web).

Information Searching Behavior is the ‘micro-level’ of behavior employed by the searcher in interacting with information systems of all kinds. It consists of all the interactions with the system, whether at the level of human computer interaction (for example, use of the mouse and clicks on links) or at the intellectual level (for example, adopting a Boolean search strategy or determining the criteria for deciding which of two books selected from adjacent places on a library shelf is most useful), which will also involve mental acts, such as judging the relevance of data or information retrieved.

Information Use Behavior consists of the physical and mental acts involved in incorporating the information found into the person’s existing knowledge base. It may involve, therefore, physical acts such as marking sections in a text to note their importance or significance, as well as mental acts that involve, for example, comparison of new information with existing knowledge. »¹⁷⁰

Esta distinção entre quatro conceitos de comportamento informacional está relacionada com a distinção que Wilson faz entre pesquisa da informação e busca da

¹⁷⁰ Cf. Wilson (2000), p. 49.

informação, a qual é visualizável no seu modelo em ninho¹⁷¹ que se reproduz a seguir na figura 9.

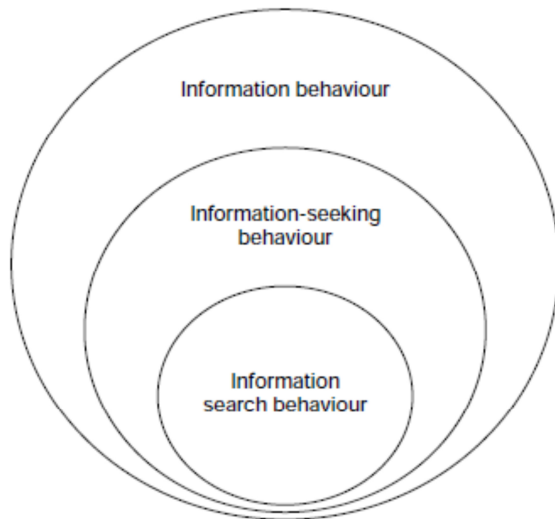


Figura 9 - Modelo em ninho das áreas de busca da informação e de pesquisa da informação de Wilson¹⁷²

Tom Wilson considera que a área do comportamento informacional é a mais abrangente, englobando o comportamento de busca da informação e este, por sua vez, o comportamento de pesquisa da informação.

Donald Case amplia, um pouco mais, o âmbito da definição de Wilson, referindo que «information behavior encompasses information seeking as well as the totality of other “unintentional” or “passive” behaviors (such as glimpsing or encountering information), as well as purposive behaviors that do not involve seeking, such as actively “avoiding” information.»¹⁷³ Nesta definição, o comportamento informacional abarca o comportamento ativo de “evitar a informação”, o que se enquadra bem nas questões relacionadas com o excesso de informação que aqui são estudadas. A ecologia da informação e a literacia da informação são dois temas do comportamento informacional que marcam presença nos currículos dos cursos de CI.

¹⁷¹ Cf. Wilson (2000), p. 63.

¹⁷² Cf. Wilson (1999), p. 263.

¹⁷³ Cf. Case (2012), p. 5.

Capítulo 3

O sistema QUETTA ou um Modelo Integrado de Estudo do Comportamento do Utilizador

3.1. Os arquivos e as bibliotecas digitais em Portugal: a constituição da amostra e a sua caracterização

O movimento *Open Access* ou “Acesso Livre ao Conhecimento” veio causar, no início do séc. XXI, o impacto inverso que a RI e a CI causaram, na segunda metade do século XX, no nicho da economia da informação onde se inserem os *information brokers*. O acesso livre ao conhecimento começou por abranger a dita “literatura cinzenta”, a produção científica de autor como as teses e dissertações e os *pre-prints* de publicações científicas, e acabou por se difundir às publicações científicas,

habitualmente, controladas por grandes grupos editoriais. Este movimento impulsionou os repositórios digitais em todo o mundo e conduziu à definição de políticas de acesso à informação no seio da comunidade científica e das instituições do ensino superior.

Historicamente, o arXiv¹⁷⁴ foi o primeiro repositório digital de acesso aberto em todo o mundo, tendo surgido em Agosto de 1991, inicialmente sob a sigla LANL, no “Los Alamos National Laboratory” nos EUA, tendo sido criado para ser um arquivo de *pre-prints* na área das ciências exatas. Paul Ginsparg¹⁷⁵ será o primeiro nome a reter na história dos repositórios digitais e, mais tarde, outros nomes do “Los Alamos National Laboratory” viriam para a ribalta como Herbert Van de Sompel e Carl Lagoze, nomeadamente, no pontapé de saída dado ao movimento *Open Archives Initiative*¹⁷⁶ (OPAI), o qual resultaria no movimento *Open Access*, com a Convenção de Santa Fé em 21-22 de Outubro de 1999¹⁷⁷. Em Dezembro de 2001, arrancou a *Budapest Open Access Initiative*¹⁷⁸ com a assinatura da Declaração de Budapeste que se tornou num documento fundamental da política de acesso livre ao conhecimento, Seguiram-se, em 2003, as outras duas declarações conhecidas, em conjunto com esta, como as 3Bs: a “Bethesda Statement on Open Access Publishing”¹⁷⁹ e a “Declaração de Berlim sobre o Acesso Livre ao Conhecimento nas Ciências e Humanidades”¹⁸⁰. Entre 10 e 12 de Dezembro de 2003, a Cimeira Mundial sobre a Sociedade da Informação, reunida em Genebra, defende, explicitamente, o acesso livre à informação científica na sua declaração de princípios e no seu plano de ação¹⁸¹. A estas declarações sucederam-se outras, logo em 2004 e nos anos seguintes, da parte da OCDE (ministros de trinta e quatro países ligados à ciência e à tecnologia), de governantes e de responsáveis por instituições do ensino superior, em diferentes localizações a nível mundial.

¹⁷⁴ <http://arxiv.org/help/general>

¹⁷⁵ http://openaccess.sdum.uminho.pt/?page_id=348

¹⁷⁶ <http://www.openarchives.org/OAI/OAI-organization.php>

¹⁷⁷ <http://www.openarchives.org/sfc/sfc.htm>

¹⁷⁸ <http://www.budapestopenaccessinitiative.org/>

¹⁷⁹ <http://legacy.earlham.edu/~peters/fos/bethesda.htm>

¹⁸⁰ <http://openaccess.mpg.de/>

¹⁸¹ http://www.itu.int/ws/is/documents/doc_multi.asp?lang=en&id=1161|1160

Em Portugal, um dos primeiros repositórios digitais a aparecer foi o RepositóriUM¹⁸² da Universidade do Minho¹⁸³ que cedo se tornou o símbolo do *Open Access* em Portugal e um projeto com visibilidade internacional. À data do seu aparecimento, em fins de Novembro de 2003, estavam registados no ROAR pouco mais de duas centenas de repositórios em todo o mundo. Em fins de 2013, já eram três mil quinhentos e dezanove (3519) no total. No sentido de selecionar a amostra de repositórios digitais a avaliar no âmbito deste projeto de investigação, foi efetuado um levantamento dos repositórios digitais existentes em Portugal em 2009. Naquela data, existiam trinta e nove (39) repositórios de Portugal registados no ROAR¹⁸⁴ e vinte e sete (27) no OpenDOAR¹⁸⁵.

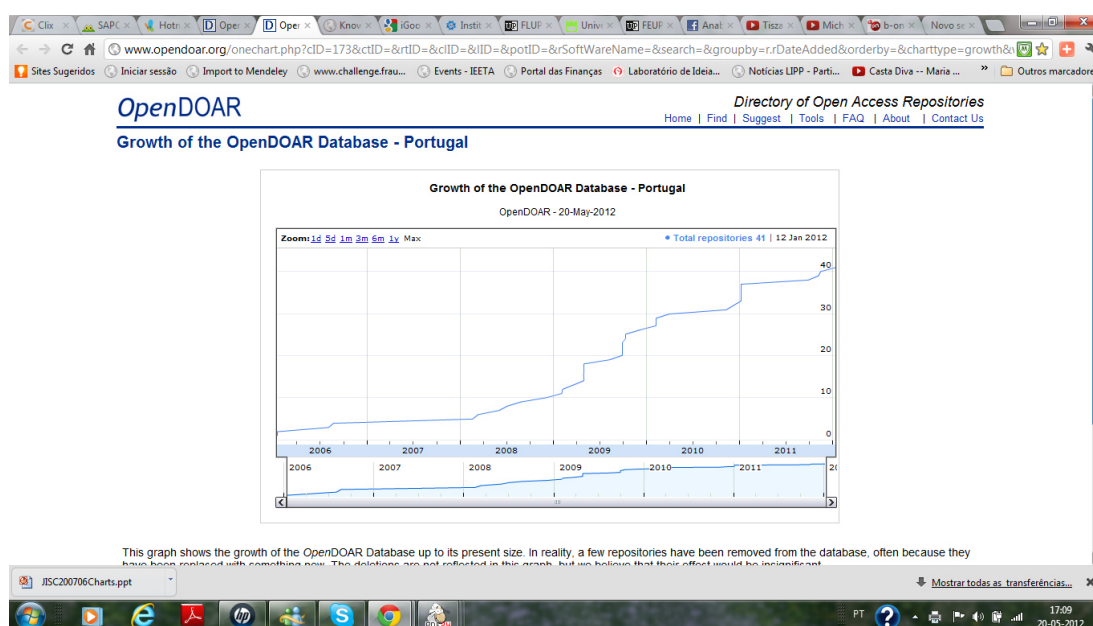


Figura 10 – Gráfico relativo à evolução dos registos de repositórios de Portugal no OpenDOAR entre 1 de Janeiro de 2006 e 12 de Janeiro de 2012

O levantamento de repositórios digitais, atrás referido, deu origem a uma listagem dos repositórios digitais sediados em Portugal em 2009 que se encontram registados

¹⁸² Apresentado em 20 de Novembro de 2003.

¹⁸³ O Departamento de Sistemas de Informação da Universidade do Minho teve um repositório designado Papadocs que foi, igualmente, um dos primeiros repositórios a serem registados no Registry of Open Access Repositories (ROAR) e que teve registo de atividade entre 8 de Março de 2004 e 2006.

¹⁸⁴ Em 2 de Dezembro de 2009.

¹⁸⁵ No OpenDOAR havia, apenas, doze (12) repositórios para Portugal em 1 de Janeiro de 2009 e em 1 de Dezembro de 2009 já eram vinte e sete (27), valor que se manteve até 1 de Janeiro de 2010.

quer no ROAR quer no OpenDOAR, a qual se encontra reproduzida na tabela 3 constante do anexo 3. A figura 10, atrás reproduzida, ilustra a evolução do número de repositórios em Portugal de acordo com o seu registo no OpenDOAR entre 1 de Janeiro de 2006 e 12 de Janeiro de 2012.

Os critérios adotados na seleção da amostra de repositórios digitais a avaliar, para cada repositório digital e para o conjunto dos repositórios selecionados, foram os seguintes:

- Maior número de registos/documentos depositados;
- Envolvimento de investigadores no desenvolvimento do repositório digital;
- Antiguidade do repositório digital;
- Atualização do conteúdo;
- Maior diversidade possível de IU entre os repositórios digitais selecionados;
- Tipologia do repositório;
- Ligação a uma instituição de ensino superior;
- Não constituir um agregador de repositórios digitais.

Face aos dados disponíveis, quer no ROAR quer no OpenDOAR, e nas restantes fontes de informação consultadas, verificou-se que os repositórios que tinham equipas de investigação dedicadas ao seu desenvolvimento e que tinham, entre outros aspetos, o maior número de documentos eram o RepositóriUM da Universidade do Minho, a SinBAD da Universidade de Aveiro, os repositórios da Universidade do Porto – o Repositório Aberto e o Repositório Temático – e a Memórias de África (designada hoje em dia por “Memórias de África e do Oriente”) mantida pela Universidade de Aveiro. Motivos pelos quais foram esses os repositórios digitais selecionados para constituírem a amostra a ser avaliada no âmbito deste projeto de investigação. Acresce referir que outro critério tido em conta foi o tipo de repositório em que se integravam – foram selecionados três repositórios institucionais e dois repositórios temáticos. Um repositório institucional e um temático da Universidade do Porto, um institucional e um temático da Universidade de Aveiro e um repositório

institucional da Universidade do Minho. Foi, ainda, considerada a diversidade das interfaces a avaliar, dado que havia um outro repositório passível de ser avaliado, quer pela importância da instituição de ensino superior a que pertence (e respetiva equipa de investigação) quer pelo número de documentos já depositados em fins de 2009, mas, que acabou por não ser considerado dada a enorme semelhança da sua IU com a do RepositoriUM da Universidade do Minho. De entre estes cinco (5) repositórios, apenas dois (2) se encontravam, então, registados no ROAR e no OpenDOAR – o RepositoriUM e o Repositório Aberto da Universidade do Porto.

Atualmente¹⁸⁶, os repositórios digitais em Portugal registados no OpenDOAR são já quarenta e cinco (45) e localizam-se em Portugal continental (43) e nos Arquipélagos dos Açores (1) e da Madeira (1).

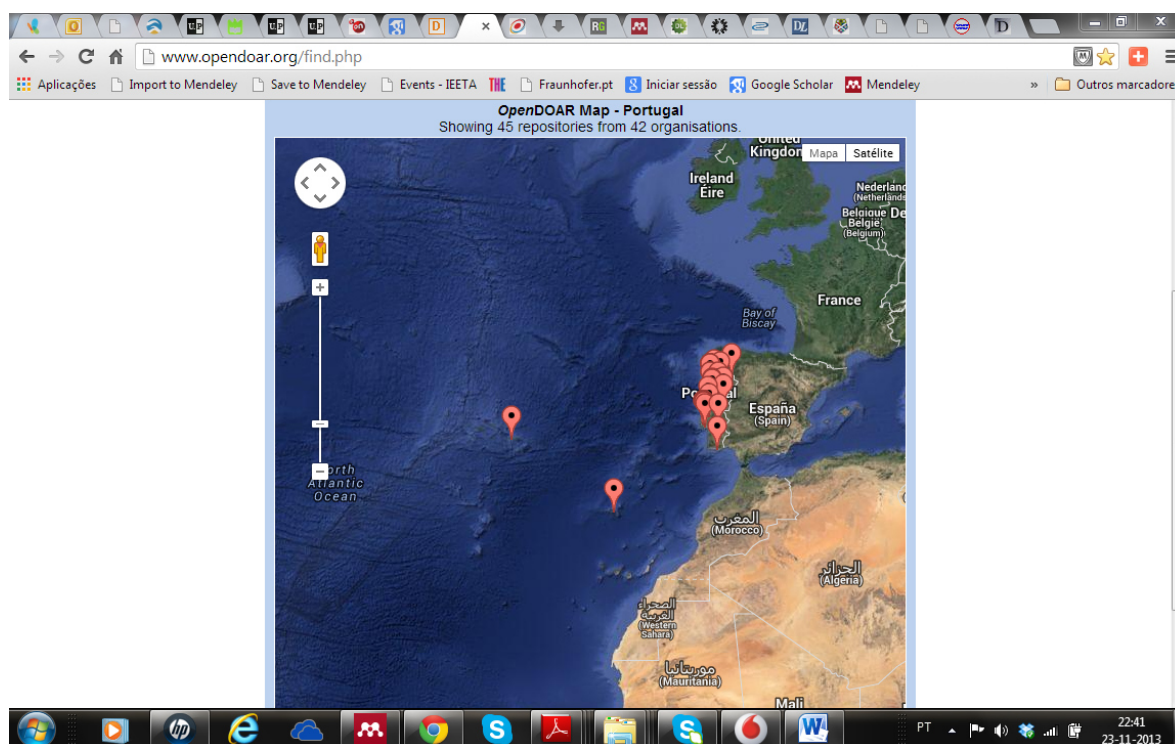


Figura 11 – Mapa dos 45 repositórios digitais existentes em Portugal continental e Arquipélagos dos Açores e da Madeira registados no OpenDOAR pertencentes a 42 organizações (23/11/2013)

¹⁸⁶ No ROAR são cinquenta e oito (58), mas verifica-se alguma duplicação de registos e constam registos de repositórios que já não se encontram disponíveis, o que os reduz a trinta e seis (36). Considerando essas imprecisões, o facto de que o OpenDOAR apresenta uma ficha de registo mais completa e de que disponibiliza ferramentas mais interessantes para a visualização dos dados, procedeu-se a um cruzamento dos dados para obter uma listagem mais fiável.

Na página anterior e nas páginas seguintes são apresentados mapas com a distribuição geográfica dos repositórios digitais atrás mencionados, de forma a elucidar acerca da localização das organizações nas quais os mesmos se encontram sediados e a partir das quais são geridos.

Os mapas foram gerados a partir das ferramentas disponíveis no *site* do OpenDOAR¹⁸⁷.

Na figura 11, atrás reproduzida, e na figura 12 que segue, pode ver-se a distribuição geográfica dos repositórios digitais em Portugal continental e ilhas adjacentes.

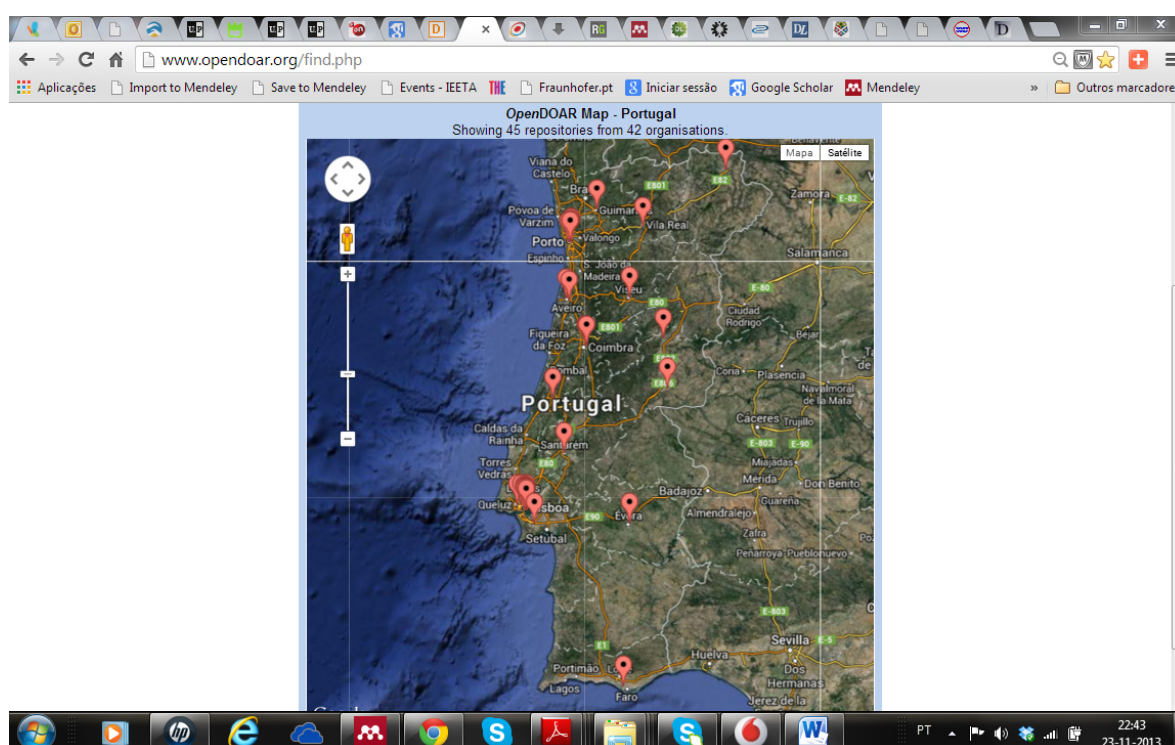


Figura 12 - Mapa dos 43 repositórios digitais existentes em Portugal continental registados no OpenDOAR (23/11/2013)

Nos arquipélagos dos Açores e da Madeira situam-se dois (2) repositórios digitais, a seguir identificados:

- ❖ Região Autónoma dos Açores (1) – Repositório da Universidade dos Açores (UAçores)
- ❖ Região Autónoma da Madeira (1) - Repositório Digital da Universidade da Madeira (UMa)

¹⁸⁷ <http://www.opendoar.org/index.html>

O considerável número de repositórios digitais existentes num espaço geográfico terrestre (uma vez que o território marítimo português é vastíssimo) tão reduzido coloca Portugal numa posição relevante a nível internacional neste âmbito.

Nas páginas seguintes, é apresentada a distribuição geográfica dos repositórios digitais sediados em Portugal continental por zonas: nas zonas Norte e Centro (figura 13), na zona do Grande Porto (figura 14), na zona Sul (figura 15), e, por fim, na zona da Grande Lisboa (figura 16).

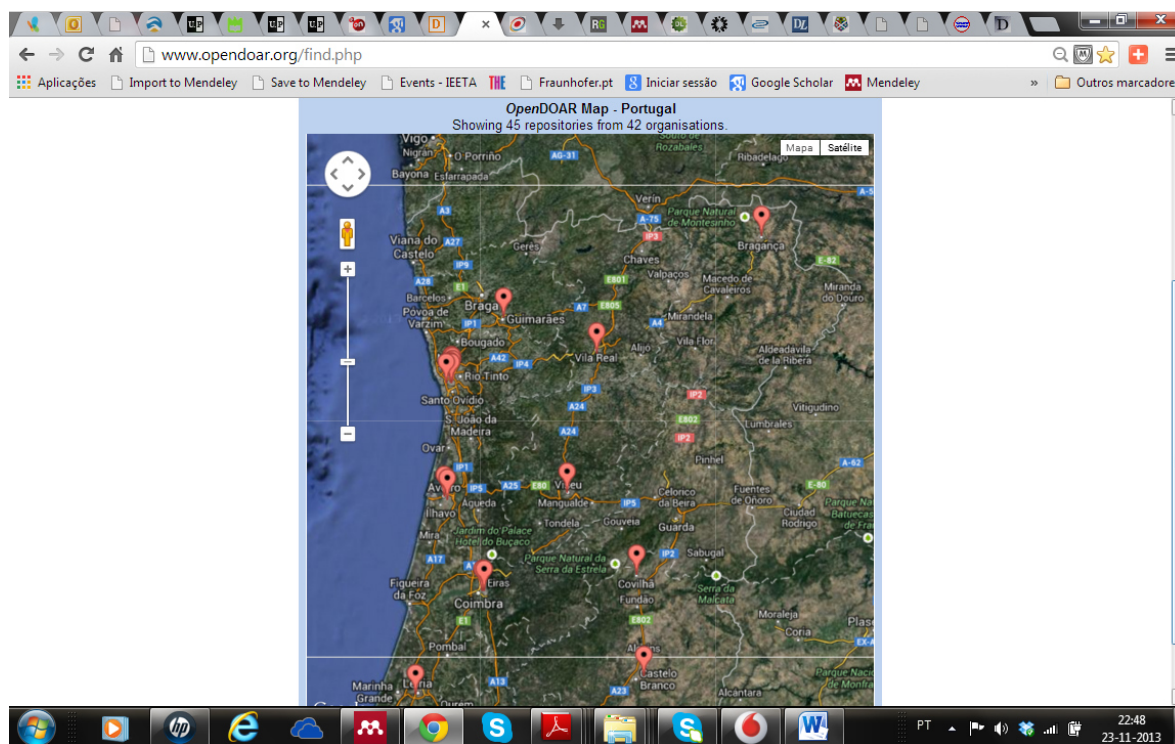


Figura 13 - Mapa dos repositórios digitais existentes nas zonas Norte e Centro do território continental de Portugal registados no OpenDOAR (23/11/2013)

Nas zonas Norte e Centro, podemos identificar vinte (20) repositórios digitais distribuídos entre o litoral e o interior do país, com maior incidência no litoral. Todos estes repositórios digitais estão sediados em instituições do ensino superior à exceção de dois que pertencem a dois centros hospitalares públicos – o Centro Hospitalar do Porto (CHP) e o Centro Hospitalar e Universitário de Coimbra (CHUC). Ressalve-se, porém, que estes dois centros hospitalares incluem dois hospitais universitários

ligados a duas instituições de ensino superior público: a Universidade do Porto e a Universidade de Coimbra. Estes vinte repositórios digitais estão, na sua totalidade, sediados em centros urbanos, mais concretamente em dez (10) cidades: Bragança, Guimarães, Vila Real, Porto, Viseu, Aveiro, Covilhã, Coimbra, Castelo Branco e Leiria.

Listam-se, abaixo, os referidos repositórios por localização geográfica, exceto os da zona do Grande Porto que serão listados mais adiante:

- ❖ Bragança (1) – Biblioteca Digital do IPB (Instituto Politécnico de Bragança)
- ❖ Guimarães (2) – e-Learning Repository (DSI da Universidade do Minho)
- RepositoriUM (Universidade do Minho)
- ❖ Vila Real (1) – Repositório da UTAD (Universidade de Trás-os-Montes e Alto Douro)
- ❖ Porto (7 repositórios listados mais adiante)
- ❖ Viseu (1) - Repositório Científico do Instituto Politécnico de Viseu (IPV)
- ❖ Aveiro (2) – LIFE+ MarPro Repository (Departamento de Biologia da Universidade de Aveiro)
- RIA Repositório Institucional da Universidade de Aveiro (UA)
- ❖ Covilhã (1) – UBI Thesis Conhecimento Online (Universidade da Beira Interior)
- ❖ Coimbra (3) – Estudo Geral (Universidade de Coimbra)
- Repositório Aberto do Instituto Superior Miguel Torga (ISMT)
- Repositório Institucional dos Hospitais da Universidade de Coimbra (CHUC)
- ❖ Castelo Branco (1) – Repositório do Instituto Politécnico de Castelo Branco (IPCB)
- ❖ Leiria (1) – ic-Online (Instituto Politécnico de Leiria)

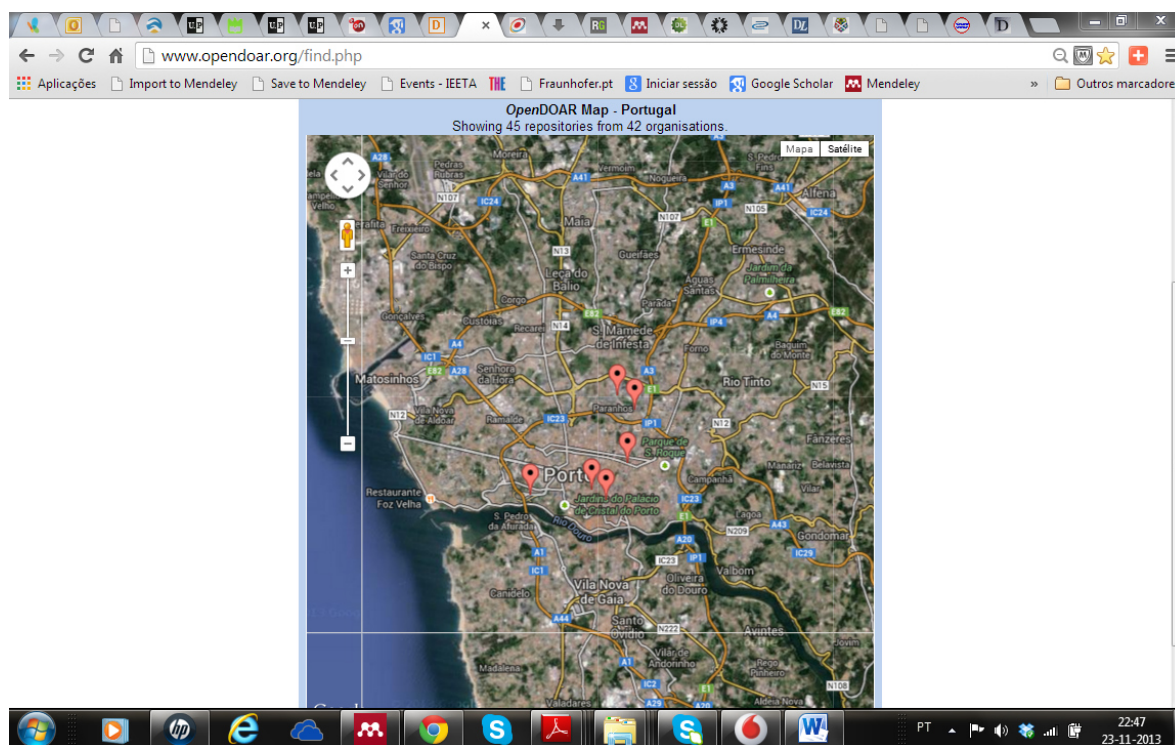


Figura 14 - Distribuição geográfica dos repositórios digitais que se localizam na zona do Grande Porto do território continental de Portugal registados no OpenDOAR (23/11/2013)

Na zona do Grande Porto, surgem sete (7) repositórios digitais, todos eles sediados na cidade do Porto. Em concreto, são os seguintes:

- ❖ Porto (7) – B-Digital (Universidade Fernando Pessoa)
 - Biblioteca Digital da FLUP (Faculdade de Letras da Universidade do Porto)
 - Repositório Aberto UPorto (Universidade do Porto)
 - Repositório Científico do Centro Hospitalar do Porto (CHP)
 - Repositório Científico do Instituto Politécnico do Porto (IPP)
 - Repositório Institucional da ESEPF (Escola Superior de Educação Paula Frassinetti)
 - Repositório Institucional da FEUP (Faculdade de Engenharia da Universidade do Porto)

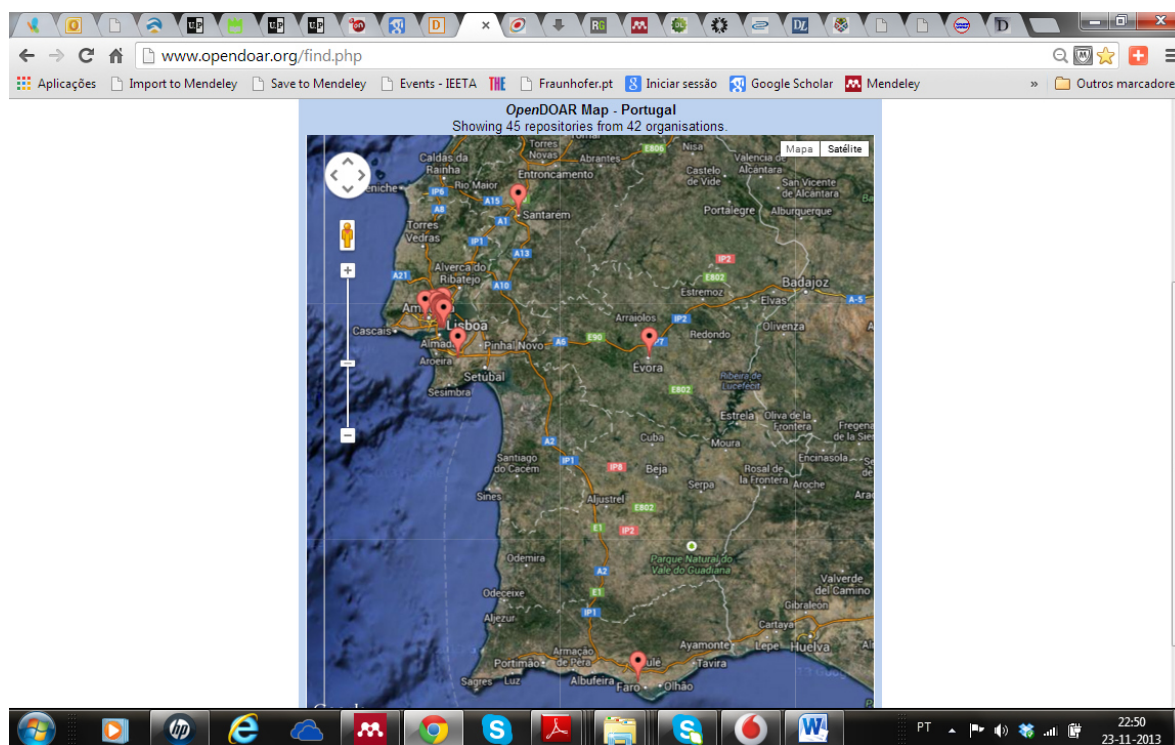


Figura 15 - Mapa dos repositórios digitais existentes na zona Sul do território continental de Portugal registados no OpenDOAR (23/11/2013)

Na zona Sul do país encontram-se sediados vinte e três (23) repositórios digitais, sendo notório o desequilíbrio entre o litoral e o interior, havendo, apenas, um repositório digital sediado no interior (na Universidade de Évora) e quatro (4) sediados fora da capital do país. Na zona Sul, excetuando os dezanove (19) da zona da Grande Lisboa, encontram-se sediados os seguintes repositórios digitais:

- ❖ Santarém (1) – Repositório Científico do Instituto Politécnico de Santarém (IPSantarém)
- ❖ Évora (1) – Repositório Científico da Universidade de Évora (UÉvora)
- ❖ Setúbal (1) – ReCiL Repositório Científico Lusófona (Grupo Lusófona)
- ❖ Faro (1) - SAPIENTIA – Repositório Institucional Universidade do Algarve (UAlg)

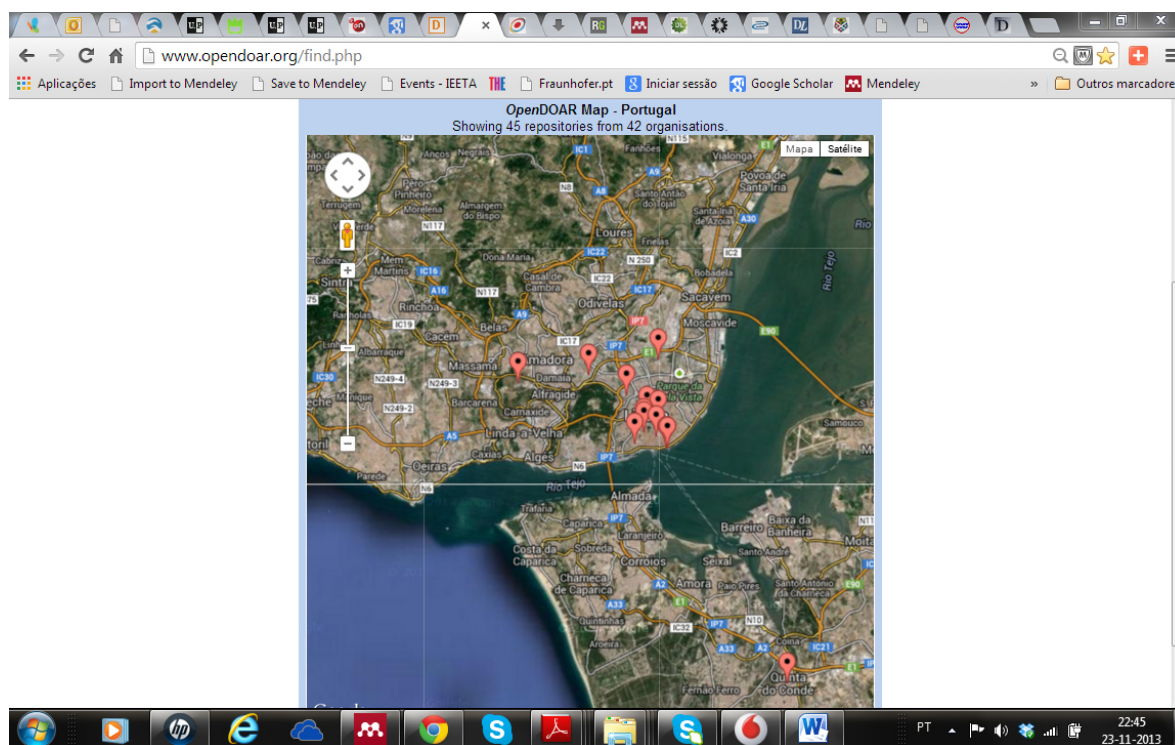


Figura 16 - Distribuição geográfica dos repositórios digitais que se localizam na zona da Grande Lisboa do território continental de Portugal registados no OpenDOAR (23/11/2013)

A zona da Grande Lisboa concentra as sedes de dezanove (19) repositórios digitais, à semelhança daquilo que sucede com as organizações em geral em Portugal, localizando-se todas na cidade de Lisboa. Como se analisou atrás, sucede o mesmo no caso da zona do Grande Porto.

Assim, na cidade de Lisboa estão sediados os seguintes repositórios digitais:

- ❖ Lisboa (19) – Repositório Científico do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA)
 - RCAAP Repositório Científico Aberto de Portugal (MCTES¹⁸⁸/UMIC¹⁸⁹)
 - Repositório Comum RCAAP (MCTES/UMIC)
 - Repositório Científico do Instituto Politécnico de Lisboa (IPL)
 - Repositório do Hospital Prof. Doutor Fernando Fonseca (HFF)
 - Repositório da Universidade Nova de Lisboa (UNL)

¹⁸⁸ Atual Ministério da Educação e Ciência (MEC).

¹⁸⁹ <http://www.unic.pt/>

- REPAP – Repositório da Administração Pública (INA)
- UTL Repository (Universidade Técnica de Lisboa)
- Repositório Aberto (Universidade Aberta)
- ARCA – IGC Repository (Instituto Gulbenkian de Ciência)
- Repositório da Universidade de Lisboa (UL) - DigiTool
- Repositório.UL (Universidade de Lisboa) – DSpace
- SciELO Portugal (MCTES, atual MEC)
- DiTeD - Depósito de Dissertações e Teses Digitais (Biblioteca Nacional)
- Repositório do ISCTE (Instituto Superior de Ciências do Trabalho e da Empresa / Instituto Universitário de Lisboa)
- Repositório do ISPA (Instituto Superior de Psicologia Aplicada)
- Repositório do LNEG (Laboratório Nacional de Energia e Geologia)
- Veritati - Repositório Institucional da Universidade Católica Portuguesa (UCP)
- Repositório do Centro Hospitalar de Lisboa Central, EPE (CHLC)

Nas figuras 17 a 22, inclusive, é apresentada a localização geográfica de cada um dos repositórios avaliados neste projeto de investigação e a sua caracterização no registo, atualmente, existente no OpenDOAR.

Nas figuras 17 e 18, surge o RepositóriUM¹⁹⁰ da Universidade do Minho, aquele que se localiza mais a norte no mapa de Portugal de entre os repositórios digitais selecionados.

¹⁹⁰ <http://repositorium.sdum.uminho.pt/>

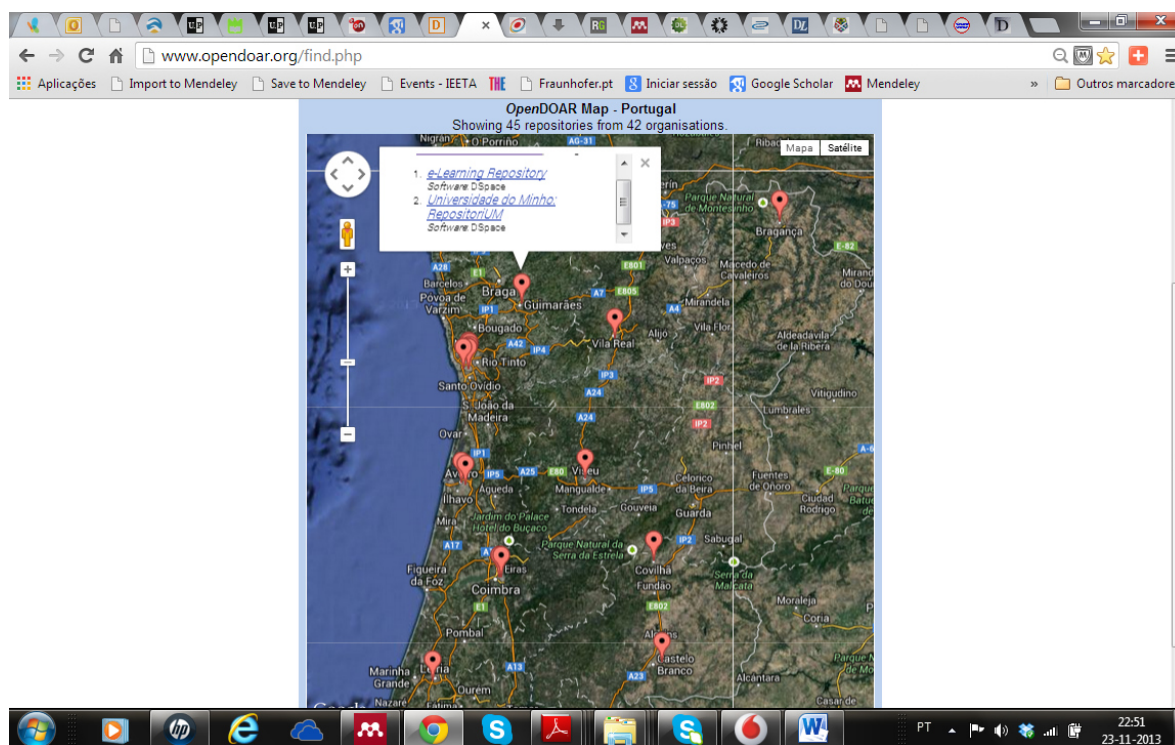


Figura 17 – Localização geográfica do RepositoriUM da Universidade do Minho na zona norte de Portugal continental registados no OpenDOAR (23/11/2013)

O **RepositoriUM** teve início em Novembro de 2003¹⁹¹, tendo a fase de testes decorrido a partir de Junho do mesmo ano, e foi desenvolvido na plataforma DSpace pela Universidade do Minho. O Reitor da Universidade do Minho assinou a Declaração de Berlim em 30 de Novembro de 2004¹⁹² e a 6 de Dezembro de 2004 definiu a política de autoarquivo da produção científica da UMinho no Despacho Reitoral RT-56/2004. Este despacho é revogado em 2010, mediante o Despacho RT-98/2010, aquando da definição de uma nova política de *Open Access*. O RepositoriUM atingiu os vinte mil (20000) documentos em Dezembro de 2010 e os dez milhões (10000000) de *downloads* em Outubro de 2013 nas vésperas do seu décimo aniversário.

¹⁹¹ <http://openaccess.sdum.uminho.pt/oatimeline/>

¹⁹² <http://openaccess.mpg.de/319790/Signatories>

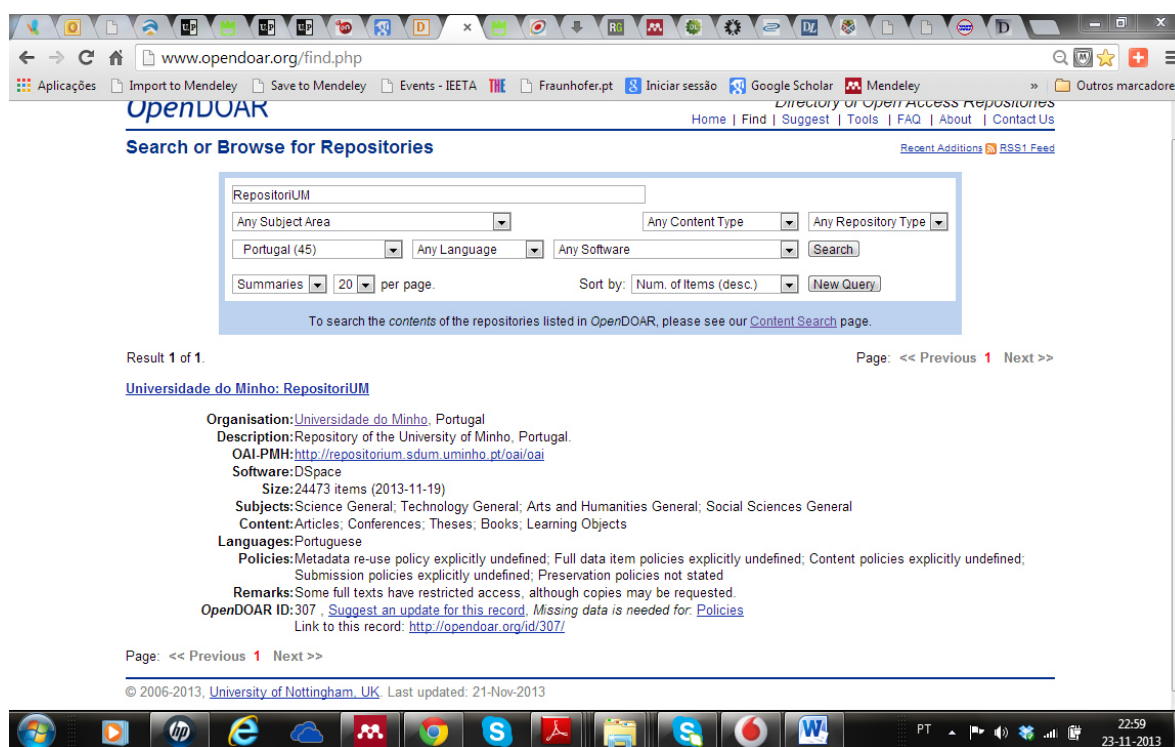


Figura 18 - Registo do RepositóriUM da Universidade do Minho no OpenDOAR (23/11/2013)

Nas figuras 19 e 20, surgem os repositórios da Universidade do Porto.

O **Repositório Aberto da Universidade do Porto**¹⁹³ foi registado no ROAR a 15 de Outubro de 2007 e foi tornado público no dia 3 de Novembro de 2007, apresentando dez mil novecentos e quarenta e seis (10946) objetos digitais a 18 de Dezembro de 2009, e atingindo trinta mil e quarenta e sete (30047) objetos digitais a 15 de Novembro de 2013. Tendo o Conselho de Reitores das Universidades Portuguesas (CRUP) aprovado e publicado a declaração do CRUP “Acesso Livre à Literatura Científica” em 28 de Novembro de 2006 e assinado a Declaração de Berlim em 4 de Janeiro de 2007, a Universidade do Porto (UP) aderiu a partir dessas datas, inerentemente, aos princípios de acesso livre à literatura científica expressos nas referidas declarações. Consequentemente, a 10 de Setembro de 2008, o Senado da Universidade do Porto aprovou a política de acesso livre (*Open Access*) da UP¹⁹⁴.

¹⁹³ <http://repositorio-aberto.up.pt/>

¹⁹⁴ http://repositorio.up.pt/files/Regulamentos_Open_Access.pdf

A plataforma selecionada foi a DSpace, tendo sido criada uma IU que permite a ligação ao SIGARRA, o SI da UP, ao ALEPH, o sistema de gestão integrada de bibliotecas utilizado na UP, e a outro *software* utilizado na UP para desenvolver repositórios digitais – o DigiTool. O Repositório Aberto era em 2009, e continua a ser em 2013, o maior repositório digital exclusivamente português¹⁹⁵, sendo, apenas, ultrapassado pela biblioteca digital Memórias de África e do Oriente¹⁹⁶ que tem colaboração de instituições de Angola e de Moçambique.

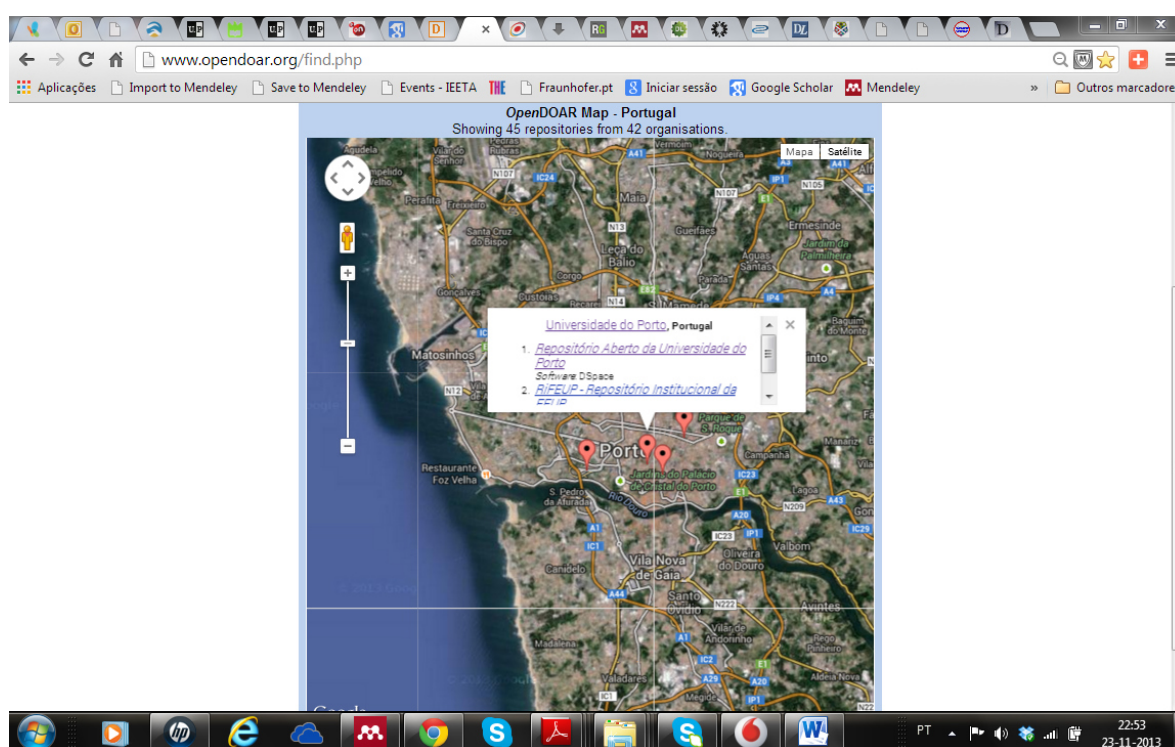


Figura 19 - Localização geográfica dos Repositórios digitais da Universidade do Porto na zona norte de Portugal continental registados no OpenDOAR (23/11/2013)

A UP tem, igualmente, o **Repositório Temático**¹⁹⁷ que não se encontra registado no ROAR nem no OpenDOAR e que, em fins de 2013, tinha já mais de trinta mil (30000) registos. Este repositório digital foi, também, um dos selecionados para ser avaliado no âmbito deste projeto de investigação, como foi referido atrás, e à data do estudo tinha já catorze mil novecentos e seis (14906) registos enquanto o Repositório

¹⁹⁵ Excluindo as plataformas de agregação de repositórios digitais.

¹⁹⁶ A Memórias de África e do Oriente tinha 371625 registos em 12 de Outubro de 2013 e 374298 registos em 22 de Dezembro de 2013.

¹⁹⁷ <http://repositorio-tematico.up.pt/>

Aberto tinha dezassete mil duzentos e oitenta e cinco (17285) registos. A plataforma é, igualmente, o DSpace, mas, a informação nele contida é de natureza diferente.

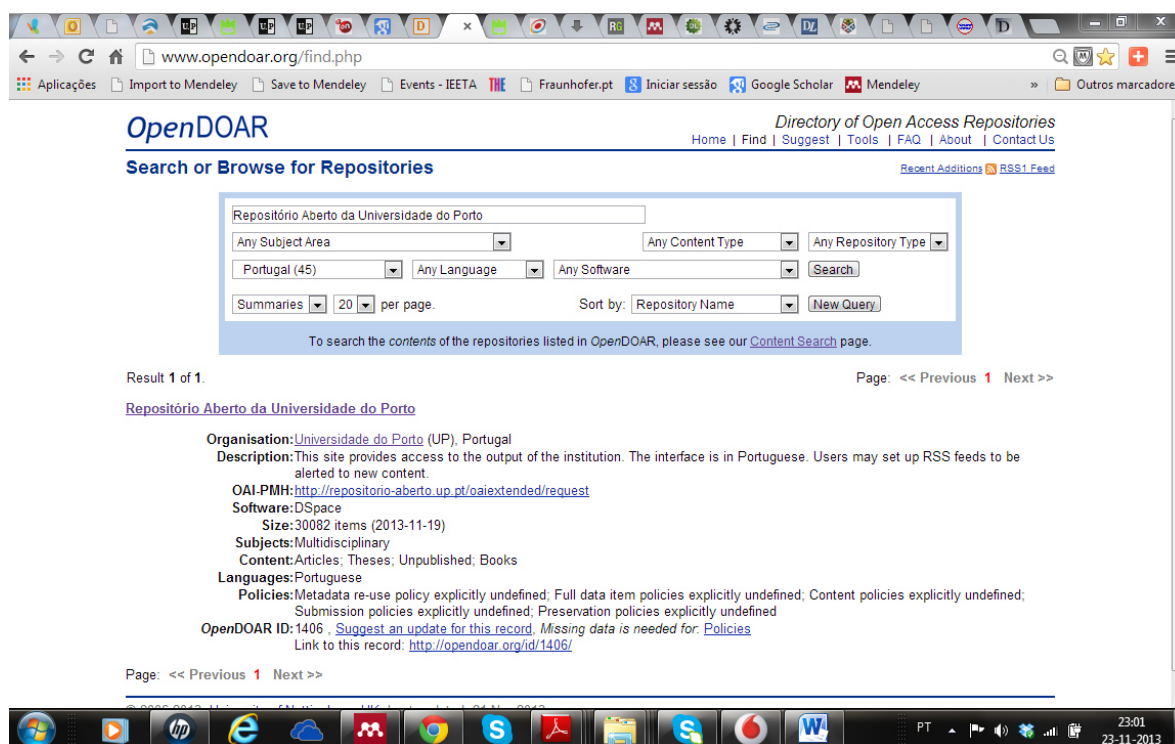


Figura 20 – Registo do Repositório Aberto da Universidade do Porto no OpenDOAR (23/11/2013)

Nas figuras 21 e 22, são apresentados os repositórios sediados na Universidade de Aveiro.

A biblioteca digital **SinBAD** foi desenvolvida na Universidade de Aveiro (UA), e implementada em 2003, tendo como característica ser uma biblioteca digital que incluía documentos de biblioteca, de arquivo e de museu¹⁹⁸. Atualmente, deu lugar a um sistema integrado de bibliotecas e arquivos digitais¹⁹⁹ e os seus registos foram migrados para as novas plataformas digitais dos Serviços de Biblioteca, Informação Documental e Museologia da UA.

¹⁹⁸ Um verdadeiro repositório digital com a abrangência da ciência da informação, pode dizer-se.

¹⁹⁹ <http://sinbad.ua.pt/>

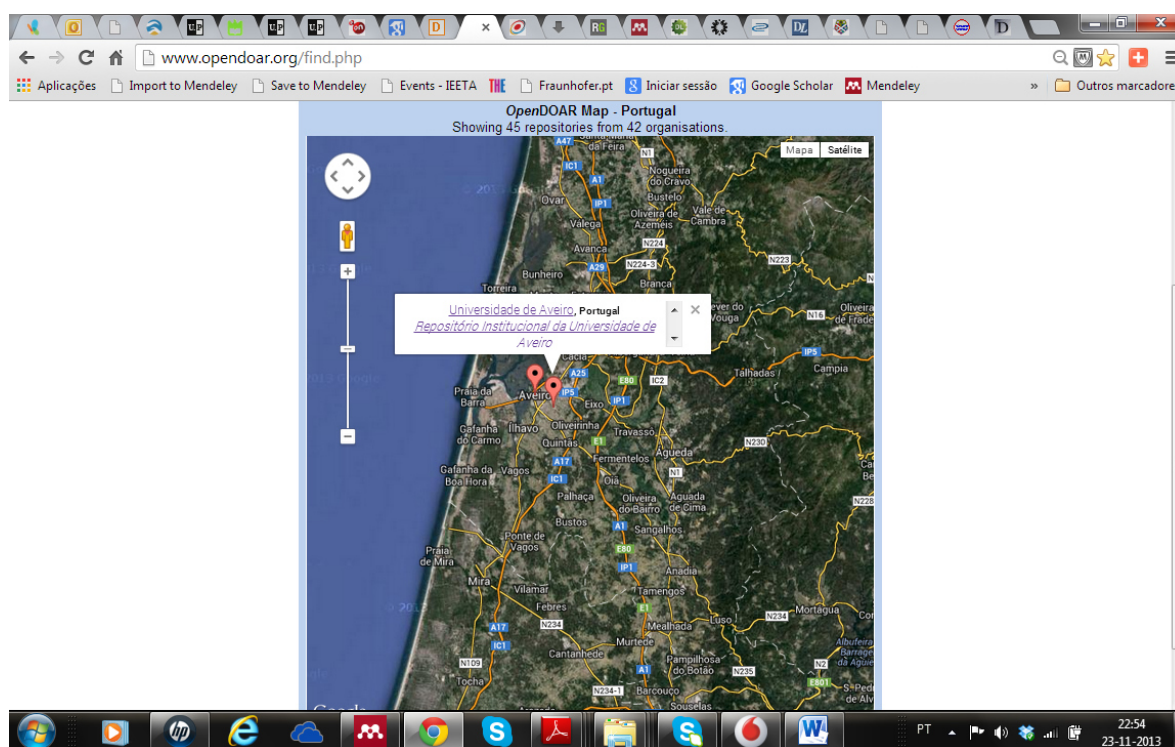


Figura 21 - Localização geográfica dos Repositórios digitais da Universidade de Aveiro na zona centro de Portugal continental registados no OpenDOAR (23/11/2013)

Os repositórios digitais indicados no mapa da figura 21 não existiam, formalmente, à data do levantamento efetuado (2009). Todavia, o RIA – Repositório Institucional da Universidade de Aveiro – integrou conteúdo da SinBAD à data do seu lançamento em fins de 2011, tendo sido migradas cerca de 3000 teses e dissertações da SinBAD para o RIA.

À data do estudo levado a cabo neste projeto de investigação (Abril/Maio de 2011), a SinBAD tinha catorze mil e duzentos e oitenta e dois (14282) documentos, quase tantos quanto tinha o Repositório Temático da UP.

A biblioteca digital **Memórias de África e do Oriente**²⁰⁰ (Memórias de África à data do estudo) integra-se num portal que constitui um projeto da Fundação Portugal-África e que foi desenvolvido e é mantido na Universidade de Aveiro, até hoje, e que teve início em 1997. Este projeto foi concretizado por uma equipa proveniente de organizações distintas: a Universidade de Aveiro (Departamento de Eletrónica e

²⁰⁰ <http://memoria-africa.ua.pt/>

Telecomunicações), o Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa (Centro de Estudos sobre África e do Desenvolvimento), o Instituto Superior de Ciências do Trabalho e da Empresa (Centro de Estudos Africanos), e o Instituto de Investigação Científica Tropical (Centro de Documentação e Informação) que coordenava o projeto. A biblioteca digital arrancou em 2005.

Em 2009, a Memórias de África tinha já cerca de duzentos e quarenta mil (240000) registos. Em Maio de 2010, tinha duzentos e sessenta e nove mil e seiscentos e oitenta e oito (269688) registos. Em fins de 2013, no dia 22 de Dezembro, a Memórias de África e do Oriente tem trezentos e setenta e quatro mil e duzentos e noventa e oito (374298) registos. Tendo em conta que o RCAAP que agrega repositórios digitais de Portugal tem, atualmente, cerca de quinhentos mil registos provenientes de sessenta e oito (68) recursos de informação, este é um volume de dados impressionante.

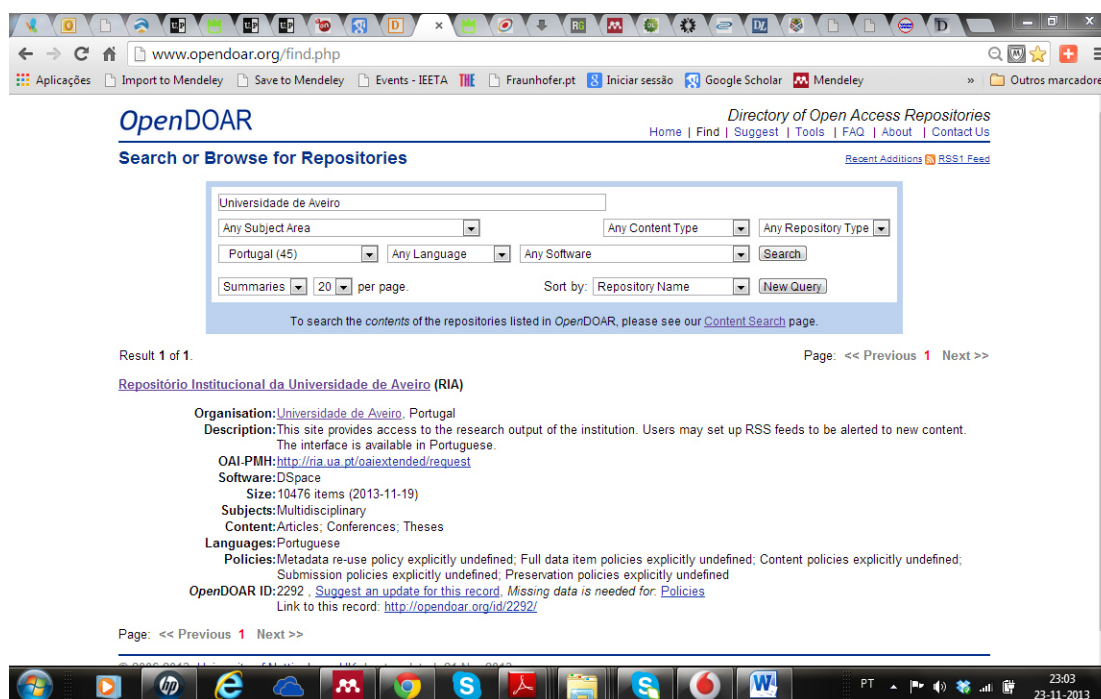


Figura 22 - Registo do Repositório Institucional da Universidade de Aveiro (RIA)²⁰¹ no OpenDOAR (23/11/2013)

²⁰¹ O RIA foi lançado em 2011, após a realização do estudo QUETTA e veio a integrar a SinBAD que foi avaliada neste estudo.

3.2. A conceção do sistema QUETTA

Após a revisão da literatura, realizada em 2008, e uma análise aprofundada das várias técnicas de investigação que poderiam ser úteis para atingir o objetivo do projeto de investigação, os métodos de investigação que pareciam poder conduzir a melhores resultados eram o experimental, o de observação e o inquérito, por motivos diversos. As técnicas de investigação utilizadas a descartar, em primeiro lugar, foram aquelas que não permitem uma contextualização na utilização dos repositórios digitais a avaliar pelos participantes no estudo – os questionários por *email* ou *online*, as entrevistas por telefone ou *chat*, o *card sort* ou *cardsorting* e os grupos de foco²⁰² (os grupos de foco são mais úteis para o desenvolvimento dos SI, pois permitem perceber o que o utilizador quer da IU do SI). Os resultados obtidos a partir destas técnicas não seriam relacionáveis com os testes de usabilidade em laboratório, porque:

- os respondentes não seriam exatamente os mesmos, e poderia não ocorrer uma única coincidência não se conseguindo que um único participante no estudo respondesse em todas as técnicas;
- o universo de respondentes de questionários por *email* seria muitíssimo alargado e não haveria a garantia de que estavam a responder após a utilização dos repositórios digitais em avaliação, e nem sequer de que, pelo menos, os conheciam e de que já os tinham pesquisado, nem de há quanto tempo não os utilizavam caso já os tivessem pesquisado.

²⁰² Esta técnica de investigação, tal como a entrevista, não permite ver com exatidão a forma como os participantes/entrevistados se comportam em interação com o produto, mas sim que racionalização fazem do seu comportamento mesmo que o façam inconscientemente, com a agravante de poderem ser influenciados pelos outros participantes que constituem o grupo. Para além disso, a análise é realizada pelo participante e não pelo observador e é assíncrona.

Os resultados finais do estudo sairiam, com grande probabilidade, enviesados, o que os tornaria pouco fiáveis, desperdiçando-se, assim, o investimento realizado no projeto de investigação. O facto de uma parte do estudo ser levada a cabo em ambiente laboratorial e de as outras componentes serem realizadas num universo mais amplo em ambiente não-controlado retiraria fiabilidade aos resultados, dado que não permitiria retirar conclusões válidas a partir do cruzamento de dados gerados em ambientes diferentes com participantes não coincidentes. Sendo possível observar os mesmos participantes em ambiente controlado ao longo de todo o estudo, nas diferentes componentes do mesmo, essa seria a situação ideal.

A observação dos participantes em ambiente controlado permite o cruzamento dos dados recolhidos, através das várias técnicas de investigação, com maior rigor, pelo que foi decidido criar um modelo integrado de estudo do comportamento do utilizador (na inexistência de um modelo que fosse adequado²⁰³) desenhado para que se conseguisse atingir o objetivo deste projeto de investigação: efetuar uma avaliação da qualidade dos SI do tipo arquivo digital e biblioteca digital, em termos de acessibilidade da informação, mais precisamente, quanto à recuperação da informação com o auxílio de SOC na IU desses repositórios digitais. A realização do estudo em ambiente controlado confere maior rigor aos resultados, permitindo captar o ponto de vista do utilizador e garantindo o controlo de todas as fases do estudo. O modelo a criar teria de integrar as várias abordagens quantitativas e qualitativas, adequadas à avaliação de bibliotecas digitais, que foram consideradas necessárias para atingir o objetivo definido - os testes de usabilidade, o inquérito, a análise dos percursos seguidos pelos utilizadores em cada sessão (*web log analysis*) e o *think-aloud*.

Mas, como se chegou a esta combinação de técnicas de investigação?

Os **testes de usabilidade** foram selecionados por permitirem obter informação pormenorizada e aquela que mais está relacionada com o utilizador real. Nos testes

²⁰³ A revisão da literatura não revelou nenhum modelo adequado ao objetivo deste projeto de investigação.

de usabilidade, o utilizador é convidado a executar tarefas reais sob observação discreta (para não condicionar a naturalidade do comportamento do utilizador) e com registo e/ou gravação dos seus comportamentos verbais e não-verbais. Permitem, ainda, perceber como o utilizador interage visualmente com a IU, e perceber se fica “perdido” na mesma. Todas estas informações permitem a posterior melhoria da IU do repositório digital quanto a usabilidade, após o diagnóstico de eventuais problemas e a redação de recomendações que conduzam à correção dos problemas encontrados.²⁰⁴

Os **testes de usabilidade com *eyetracking*** permitem, adicionalmente, o registo e a posterior análise dos movimentos oculares do utilizador no monitor do *eyetracker* aquando da visualização do SI a ser testado, ou seja, permitem a análise dos percursos seguidos pelos utilizadores em cada sessão (***web log analysis***), para além da análise das fixações em pontos localizados da *webpage* e da duração da permanência num dado ponto da *webpage* ou numa dada *webpage*. Por este motivo, esta técnica de investigação pareceu mais adequada ao perfil do estudo a realizar, já que permitia ter duas abordagens numa só e em ambiente controlado, com maior precisão para a análise e com menor cansaço para os participantes²⁰⁵. Dado que a localização dos conteúdos numa *webpage* poder ser determinante para a acessibilidade dos mesmos, optou-se pelos testes de usabilidade com *eyetracking* como técnica de investigação ideal, não obstante o carácter dispendioso do equipamento necessário para o efeito, o que poderia ser um obstáculo à realização do trabalho laboratorial. Porém, na altura de planear o trabalho laboratorial, verificou-se que não estava disponível um laboratório equipado com *eyetracker* que permitisse o cumprimento da calendarização prevista no plano do projeto de investigação (o plano de doutoramento), pelo que se chegou a ponderar a substituição deste método por estudos de usabilidade *online* de larga escala. Este tipo de estudo de usabilidade permitiria abranger uma larga amostra populacional, mas traria o inconveniente de não permitir a observação direta de cada um dos participantes, ainda que pudesse ser

²⁰⁴ Cf. J.F. Dumas, J.C. Redish (1993).

²⁰⁵ Aqui a vantagem está, claramente, na possibilidade de cruzar os dados dos testes de usabilidade com os da *web log analysis* sem descontextualização dos mesmos e com coincidência total dos participantes no estudo.

possível cruzar os dados do estudo com os do inquérito e, eventualmente, com os da *web log analysis*. Enquanto se preparava esta outra alternativa, surgiu a possibilidade de conseguir um *eyetracker* que tinha sido cedido pela Portugal Telecom SGPS SA (PT)²⁰⁶ ao polo de Aveiro do Centro de Estudos das Tecnologias e Ciências da Comunicação (CETAC.MEDIA).²⁰⁷ A solicitação da cedência do *eyetracker* a este projeto de investigação foi feita pelo Instituto de Engenharia Eletrónica e Telemática de Aveiro (IEETA)²⁰⁸, a cuja equipa a autora desta tese pertence como membro integrado. Os testes de usabilidade com *eyetracking*, tal como a *web log analysis*, implicam a recolha ou análise de enormes quantidades de dados e uma análise demorada, mesmo que a recolha de dados seja efetuada num curto intervalo de tempo.

Perante esta reviravolta nos recursos materiais disponíveis para o projeto de investigação, abandonou-se a alternativa dos estudos de usabilidade *online* de larga escala, e retomou-se a ideia de conceber um modelo integrado de estudo do comportamento do utilizador. O desenvolvimento inicial do modelo que tinha os testes de usabilidade com *eyetracking* como uma das componentes foi repescado, combinando-se aqueles com o protocolo *think-aloud* e com questionários (um questionário prévio e um questionário posterior a cada teste de usabilidade com *eyetracking*).

O protocolo *think-aloud* foi selecionado porque permite que o observador capte, de forma síncrona, as reações do participante no estudo (irritação, hesitação, confusão, etc.), e identifique e registe eventuais erros ou conceitos subjacentes. O *think-aloud* foi escolhido em substituição de uma entrevista final, a ser realizada após os testes de usabilidade, que tinha sido planeada na combinação inicial de técnicas de investigação. Tomou-se essa decisão pelo facto de a análise na entrevista ser assíncrona e ser realizada pelo participante e não pelo observador, enquanto a mesma é síncrona quer no *think-aloud* quer nos testes de usabilidade, e pelo facto de aquela

²⁰⁶ Website da PT: <http://www.telecom.pt/InternetResource/PTSite/PT/>

²⁰⁷ Website do CETAC.MEDIA: <http://www.cetacmedia.org/>

²⁰⁸ Website do IEETA: http://wiki.ieeta.pt/wiki/index.php/Main_Page

não permitir perceber a forma como o participante/entrevistado se comportou ao interagir com a IU do repositório digital, mas, antes, qual a racionalização que aquele fez do seu comportamento (ainda que o faça inconscientemente).

Os **questionários** permitem a recolha de uma grande quantidade de dados de uma forma pouco dispendiosa e abrangem amostras populacionais que podem ser bastante alargadas. Este último aspeto não foi determinante na seleção para as componentes do modelo a ser construído, dado que o estudo se destinava a ser realizado em ambiente controlado. E permitem, igualmente, caracterizar os participantes no estudo e recolher, de forma rápida, dados acerca das suas impressões imediatas sobre o repositório digital pesquisado durante o teste de usabilidade com *eyetracking*.

Na figura 23, abaixo reproduzida, pode ver-se uma distribuição de vários métodos de investigação de acordo com vários aspetos:

- por tipo de fonte dos dados;
- por tipo de abordagem;
- por contexto de utilização do produto.

Percebe-se, rapidamente, que os testes de usabilidade com *eyetracking* constituem a abordagem com maior incidência comportamental quanto aos dados recolhidos, com uma abordagem mista qualitativa e quantitativa, e com um contexto misto de utilização do produto (natural e controlado ou laboratorial). Esta constatação é suficiente para eleger esta técnica como sendo a componente preferencial no modelo em elaboração, embora não a única.

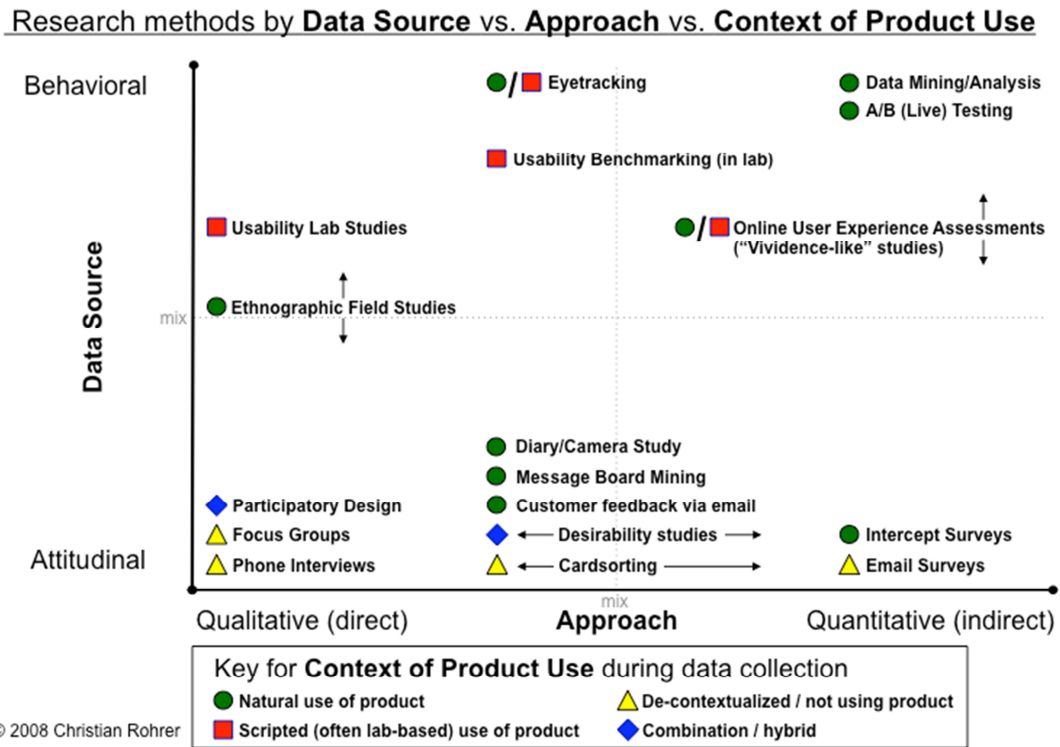


Figura 23 – Quadro sobre métodos de investigação por fonte dos dados versus Tipo de Abordagem versus Contexto de Uso do Produto reproduzido a partir do site de Jakob Nielsen²⁰⁹

Este quadro é muito elucidativo acerca da validade das escolhas feitas para a combinação de técnicas de investigação a utilizar no modelo integrado de estudo do comportamento do utilizador. Resultando da combinação de questionários, testes de usabilidade com *eyetracking*, e *think-aloud*, o modelo em questão foi batizado com a designação “sistema QUETTA”, resultando a sua sigla da extração de letras das várias técnicas de investigação componentes do modelo como a seguir se demonstra: QUestionários + testes de usabilidade com *EyeTracking* + *Think-Aloud* = QUETTA.

O sistema QUETTA – modelo integrado de estudo do comportamento do utilizador – encontra-se explicado, visualmente, nos diagramas que constam das figuras 24 e 25 inseridas a seguir à próxima página e no anexo 4, cujos passos serão, textualmente, descritos a seguir. A figura 24 apresenta o modelo passo a passo e a

²⁰⁹ <http://www.nngroup.com/articles/which-ux-research-methods/>

figura 25 apresenta o modelo com a iniciação de um contador o que facilita a sua aplicação a qualquer bateria de testes que se queira fazer numa dada investigação. O sistema QUETTA combina três técnicas de investigação em que a componente central pertence ao método experimental – os testes de usabilidade com *eyetracking* - e se alicerça no inquérito – os questionários - e no método de observação - o *think-aloud*. O estudo tem início com um questionário prévio à “bateria de testes” que se destina a caracterizar cada participante, seguindo-se os testes de usabilidade com *eyetracking* (cinco neste estudo) intervalados por um questionário posterior que é preenchido, imediatamente, a seguir ao final de cada teste e que se destina a recolher dados acerca das primeiras impressões do utilizador sobre o teste que acabou de realizar. Em simultâneo, o observador analisa discretamente, tomando notas e cronometrando²¹⁰, o comportamento do utilizador ao longo do estudo (*think-aloud*). Foi equacionada a utilização de uma câmara de vídeo para registar as sessões, para além dos registos do *eyetracker*, mas, após alguma ponderação acerca de questões relacionadas com a sua montagem e o seu posicionamento na sala do laboratório, concluiu-se que não ficaria montada nas condições ideais, o que tornaria a sua presença algo “invasiva” podendo afetar a naturalidade do comportamento dos participantes no estudo.

²¹⁰ Os dados de cronometragem manual destinaram-se a ser cruzados com os dos registos de vídeo gerados pelo *software* do *eyetracker* - Tobii Studio 2.0.5

Anabela Prista Saraiva Serrano
 Tese de Doutoramento em Informação e Comunicação em Plataformas Digitais
 Universidade do Porto / Universidade de Aveiro

Modelo do Sistema QUETTA

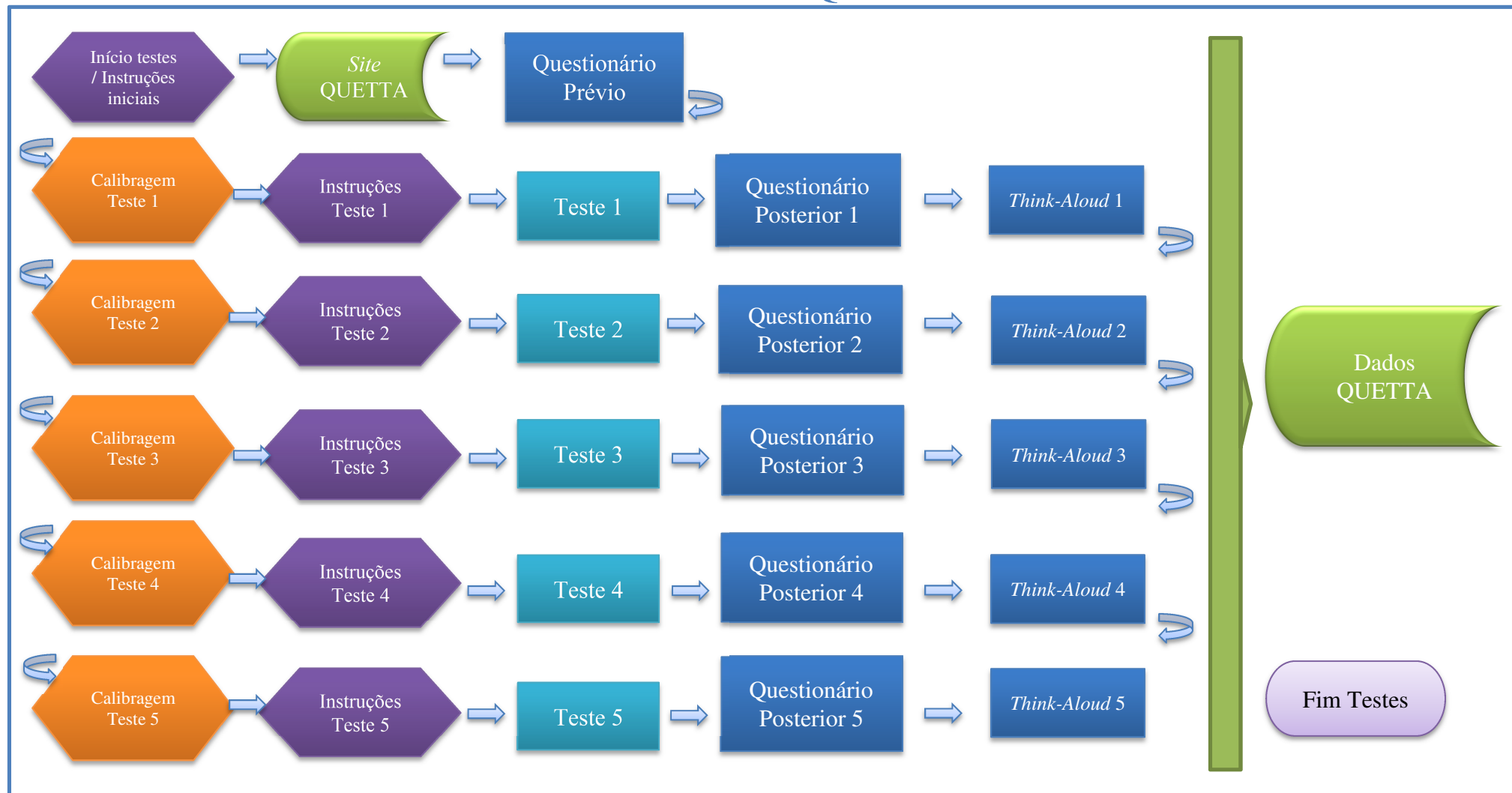


Figura 24 – Modelo do Sistema QUETTA

Modelo do Sistema QUETTA com contador

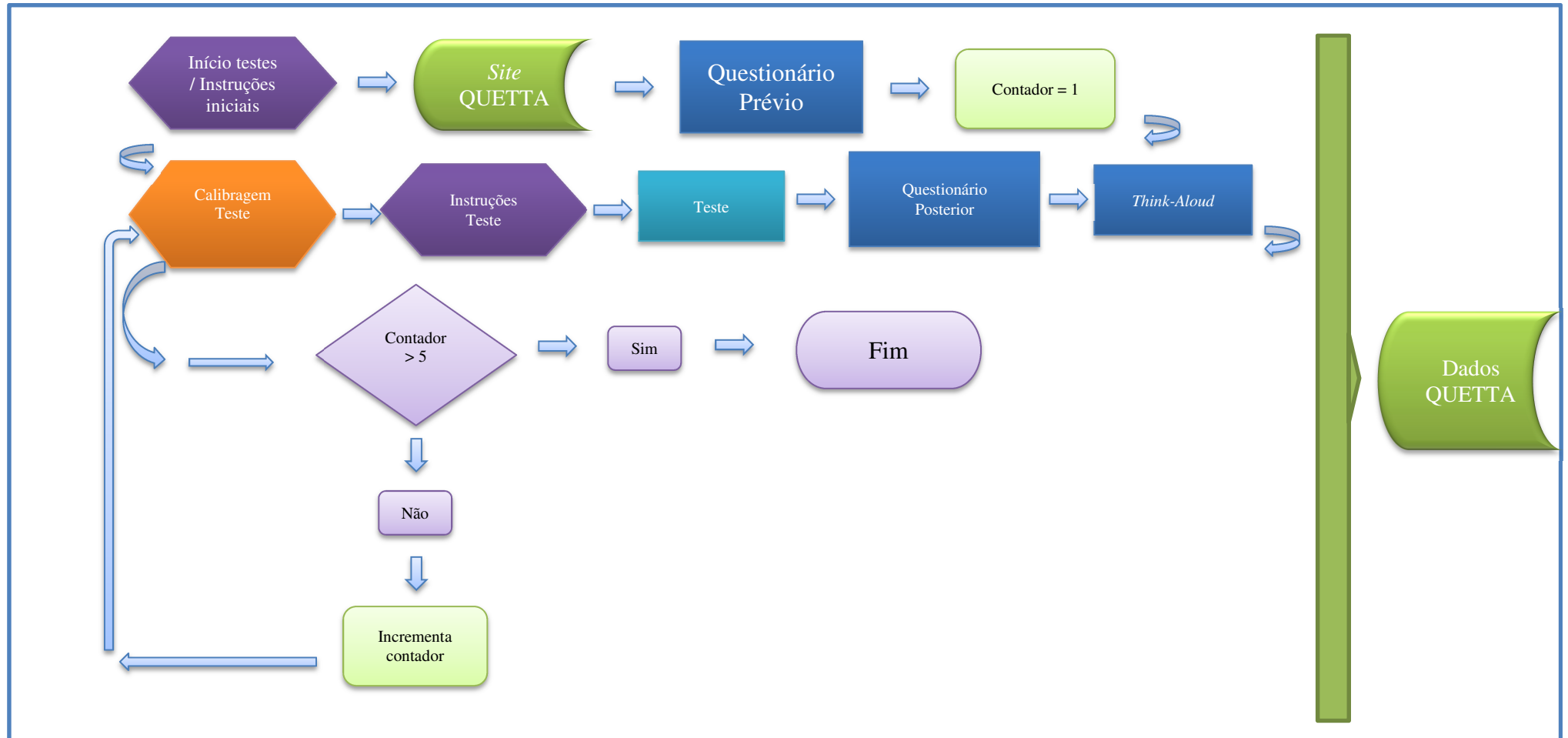


Figura 25 - Modelo do Sistema QUETTA iniciando um contador

O sistema QUETTA foi gerido através de um *website*²¹¹ criado, propositadamente, para a realização do estudo do comportamento do utilizador conduzido no âmbito deste projeto de investigação, daqui em diante designado por “*site* QUETTA”. O *site* QUETTA foi criado na plataforma da empresa Google “Google Sites”, sendo os questionários gravados noutra plataforma da empresa Google - o “Google Docs”.

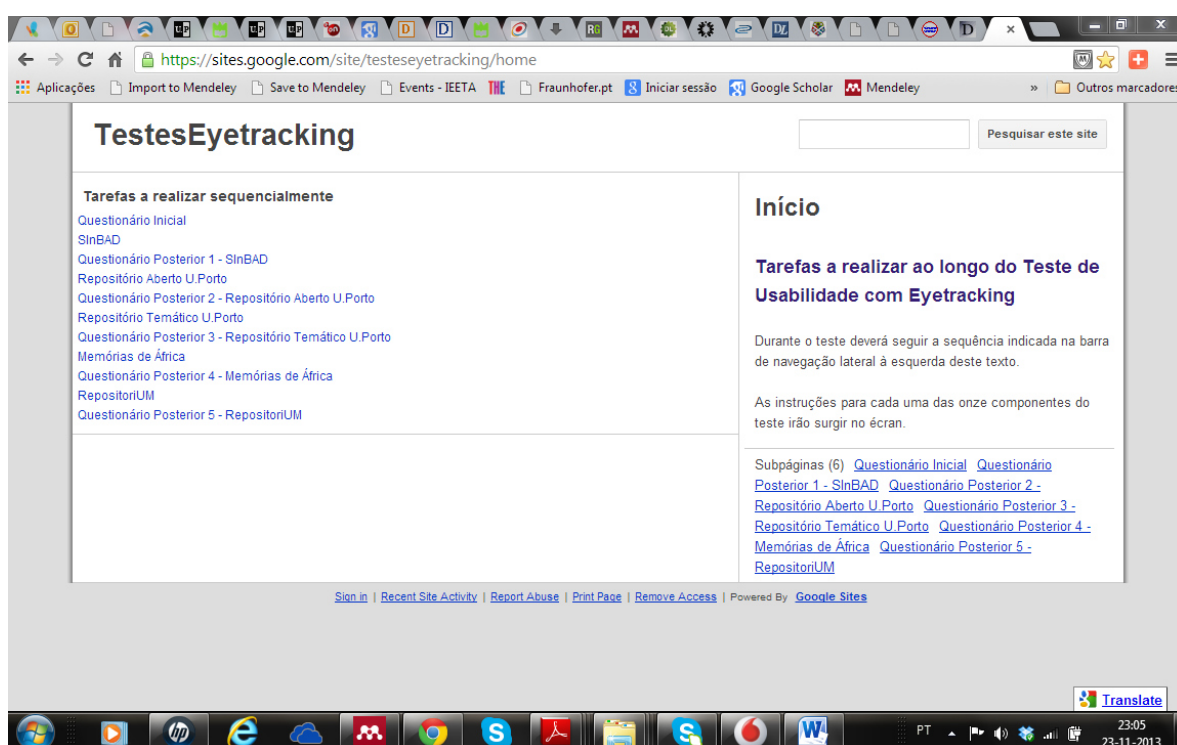


Figura 26 – Página inicial do *Site* QUETTA

Adicionalmente, foram criadas páginas com instruções prévias a cada teste de usabilidade com *eyetracking*, onde são descritas as tarefas a serem executadas ao longo do mesmo, no *software* do *eyetracker* da marca Tobii, o Tobii Studio 2.0.5, que replicavam as instruções constantes das páginas *web* do *site* QUETTA relembrando-as após a calibragem efetuada antes de cada teste ser iniciado.

²¹¹ <https://sites.google.com/site/testeseyetracking/home>

O estudo envolveu vinte e oito (28) passos em que onze (11) deles foram concretizados pelo participante e dezassete (17) foram conduzidos pelo observador. Cada participante cumpriu a sequência de passos que integram o sistema QUETTA e que se encontra descrita no protocolo do sistema QUETTA nas linhas seguintes.

Protocolo do sistema QUETTA:

1. Para iniciar o estudo, são dadas **instruções iniciais** gerais acerca da bateria de testes a realizar e dos procedimentos que irão ser executados, bem como de alguns pormenores técnicos relativos a postura e acuidade visual que podem influir nos resultados dos testes;
2. Em seguida, apresenta-se a página inicial do *site* **QUETTA** que consta na figura 26 e explica-se que para realizar os testes basta seguir as instruções do *site* e do facilitador do estudo que estará acessível a alguma distância para não perturbar a realização dos testes nem influir no resultado dos mesmos e do estudo em geral;
3. Dá-se início ao estudo, e o participante clica na ligação para a primeira tarefa a realizar, sequencialmente, na página inicial do *site* QUETTA – o questionário inicial (ou prévio) – como se pode verificar na figura 27;

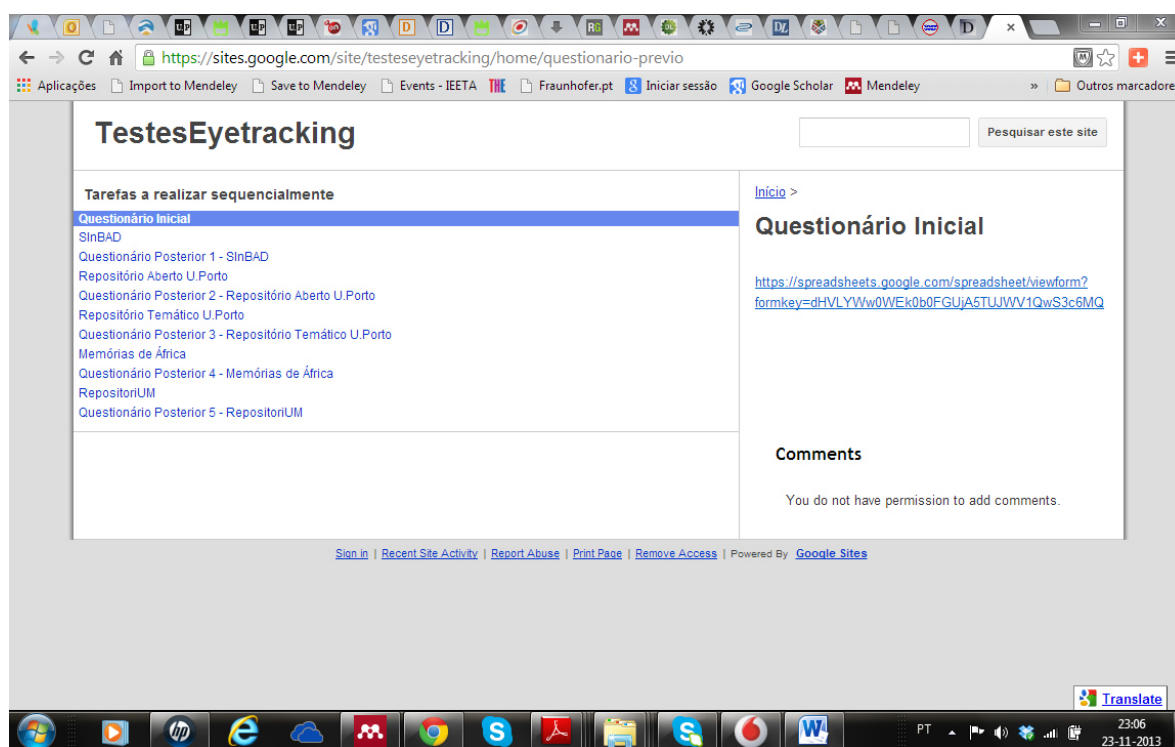


Figura 27 – Página de ligação para o questionário inicial no *site* QUETTA

4. É aberta a *webpage* onde consta o **questionário prévio** ao teste de usabilidade, como se pode ver nas figuras 28, 29, 30, 31 e 32, que é executado apenas no início do conjunto de testes e não antes de cada um deles, daí que seja designado como questionário inicial no *site* QUETTA, e que deve ser preenchido *online* pelo participante com o número de identificação do utilizador (ID) indicado pelo facilitador, o qual deverá ser usado sempre que o participante preencher questionários ao longo do estudo;
5. Acabado o preenchimento do questionário inicial no *site* QUETTA, é iniciada a **calibragem para o Teste 1** com o *eyetracker* lembrando as instruções iniciais relativas a pormenores técnicos ao participante - se a calibragem não for bem-sucedida, repete-se a calibragem²¹² descansando o participante quanto a esse procedimento;

²¹² As calibrações de má qualidade ou de qualidade sofrível inviabilizam a recolha de dados, pelo que devem ser descartadas. Se não se conseguir uma boa calibragem, em vez de se abortar o estudo, prossegue-se e marca-se esse ID como sendo para descartar no final.

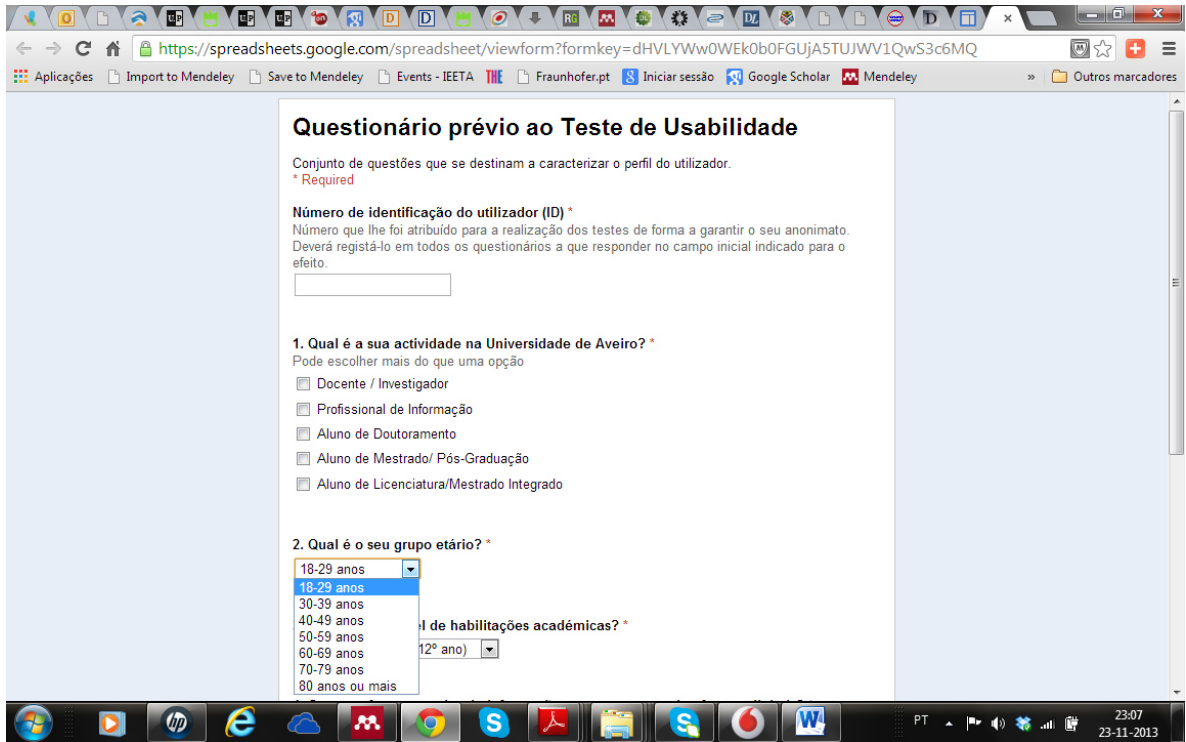


Figura 28 - Página do questionário inicial, ou questionário prévio ao teste de usabilidade, no site QUETTA (parte 1)

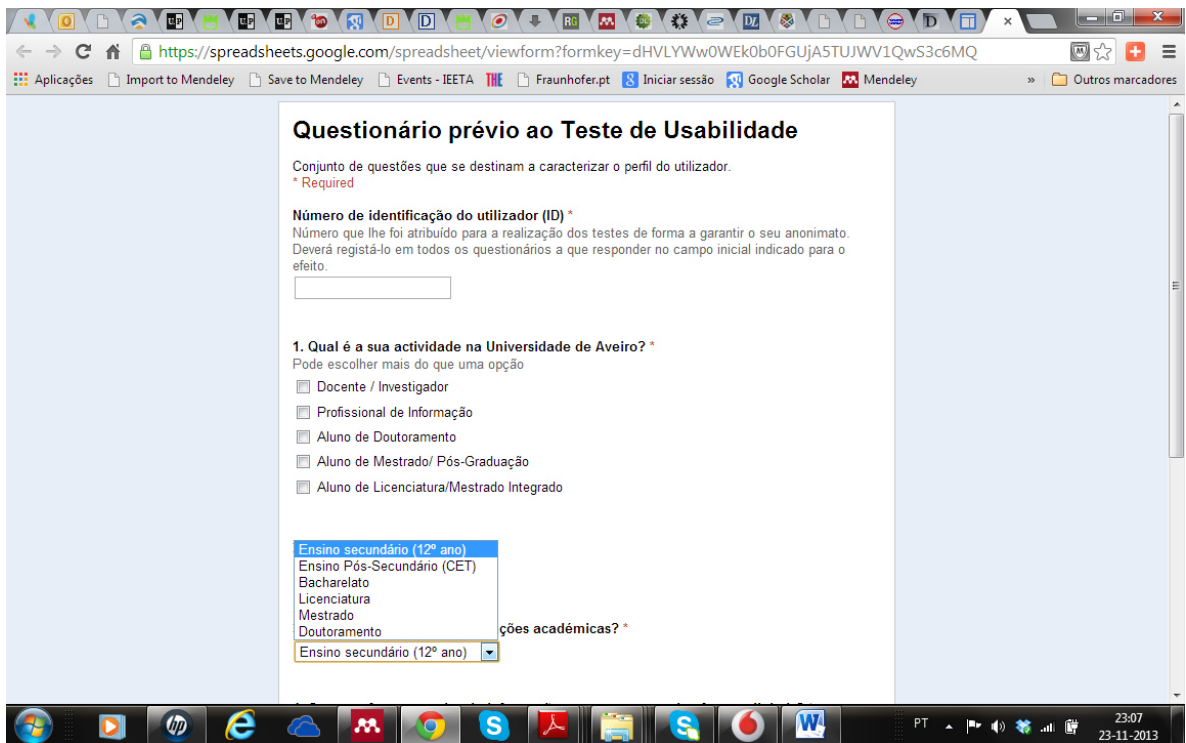


Figura 29 - Página do questionário inicial, ou questionário prévio ao teste de usabilidade, no site QUETTA (parte 2)

2. Qual é o seu grupo etário? *

18-29 anos

3. Qual é o seu nível de habilitações académicas? *

Ensino secundário (12º ano)

4. Costuma fazer pesquisa de informação por assunto em plataformas digitais? *
(Bibliotecas digitais, repositórios institucionais, arquivos digitais) Se respondeu "não" o questionário prévio termina aqui.

Sim

5. Qual a frequência com que acede a esses recursos de informação?

Frequentemente

6. A pesquisa por assunto é para si mais usual do que a pesquisa por autor e título?

Sim

Submit

Never submit passwords through Google Forms.

Powered by [Google Docs](#)

[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

Figura 30 - Página do questionário inicial, ou questionário prévio ao teste de usabilidade, no *site* QUETTA (parte 3)

2. Qual é o seu grupo etário? *

18-29 anos

3. Qual é o seu nível de habilitações académicas? *

Ensino secundário (12º ano)

4. Costuma fazer pesquisa de informação por assunto em plataformas digitais? *
(Bibliotecas digitais, repositórios institucionais, arquivos digitais) Se respondeu "não" o questionário prévio termina aqui.

Sim

5. Qual a frequência com que acede a esses recursos de informação?

Frequentemente

6. A pesquisa por assunto é para si mais usual do que a pesquisa por autor e título?

Sim

Submit

Never submit passwords through Google Forms.

Powered by [Google Docs](#)

[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

Figura 31 - Página do questionário inicial, ou questionário prévio ao teste de usabilidade, no *site* QUETTA (parte 4)

2. Qual é o seu grupo etário? *

18-29 anos

3. Qual é o seu nível de habilitações académicas? *

Ensino secundário (12º ano)

4. Costuma fazer pesquisa de informação por assunto em plataformas digitais? *
(Bibliotecas digitais, repositórios institucionais, arquivos digitais) Se respondeu "não" o questionário prévio termina aqui.

Sim

5. Qual a frequência com que acede a esses recursos de informação?

Frequentemente

6. A pesquisa por assunto é para si mais usual do que a pesquisa por autor e título?

Sim
Sim
Não

Submit

Never submit passwords through Google Forms.

Powered by [Google Docs](#)

[Report Abuse](#) - [Terms of Service](#) - [Additional Terms](#)

Figura 32 - Página do questionário inicial, ou questionário prévio ao teste de usabilidade, no *site* QUETTA (parte 5 e última)

6. Após a calibragem, dá-se início ao teste 1 com as **instruções para o teste 1** na IU do *eyetracker* como se pode verificar na figura 33;
7. O **teste 1** tem início após a leitura das instruções com a ligação ao site do repositório digital a ser avaliado: a biblioteca digital SinBAD da Universidade de Aveiro como se pode ver na figura 34;
8. Após o teste 1, o participante clica na ligação seguinte na página inicial do *site* QUETTA (cf. figura 35) que o conduz ao **questionário posterior 1** que deverá ser preenchido no final do teste 1 e que consta nas figuras 36, 37, 38 e 39;

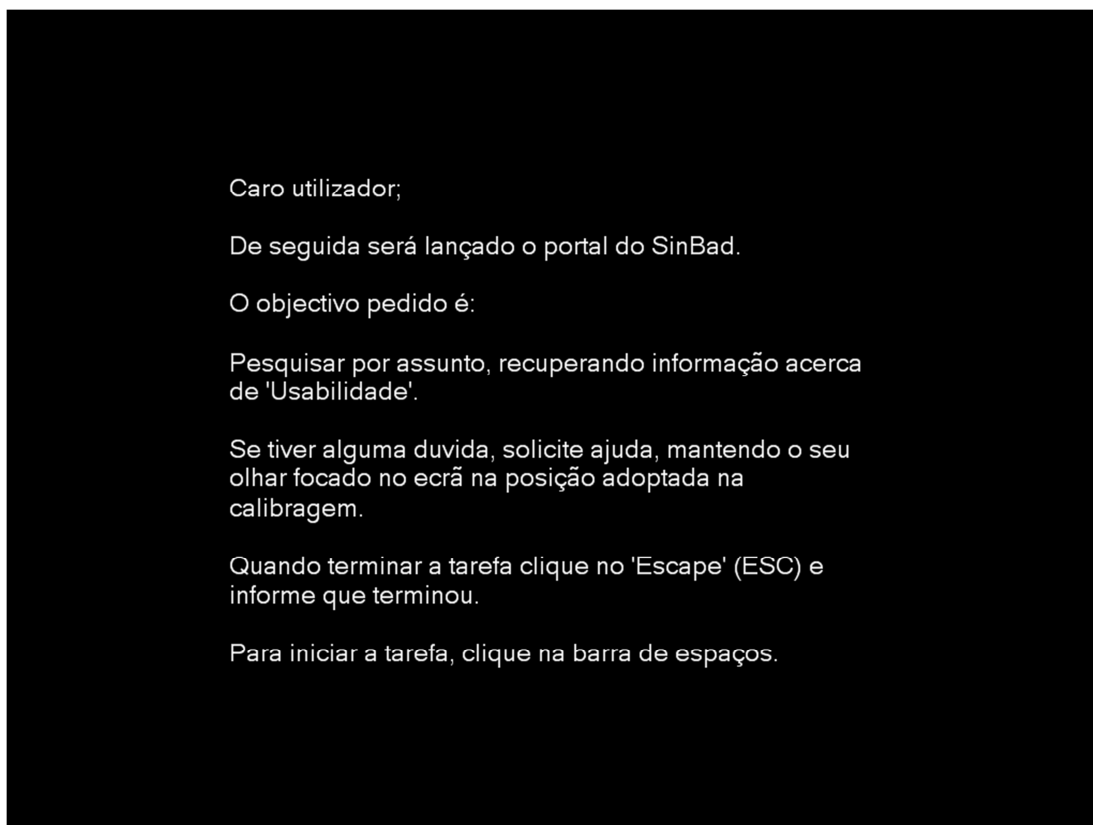


Figura 33 - Instruções para o teste 1 na IU do eyetracker



Figura 34 – Biblioteca digital SinBAD da Universidade de Aveiro (teste 1)



Figura 35 - Página de ligação para o “questionário posterior 1 – SinBAD” no site QUETTA

The screenshot shows a Google Forms interface in a browser window. The title is "Questionário posterior ao Teste de Usabilidade 1". Below the title, there is a required field for "Recurso de Informação - SinBAD" with the URL <http://biblioteca.sinbad.ua.pt/Pesquisa/?sub=4>. This is followed by a required field for "Número de identificação do utilizador (ID)", with a text input box and a description: "Número que lhe foi atribuído para a realização dos testes de forma a garantir o seu anonimato. Deverá registá-lo em todos os questionários a que responder no campo inicial indicado para o efeito." Below this are three questions:

1. Considera que foi fácil aceder à informação pretendida? *
A dropdown menu is open, showing options: Sim (selected), Não.
2. Encontrou auxílio para a pesquisa por assunto na interface de pesquisa? *
A dropdown menu is open, showing options: Sim (selected), Não.
3. Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa? *
Description: "Vocabulários que permitem identificar relações semânticas ou de significado ou relações sintáticas ou de frase entre os termos - dicionários, roteiros e redes semânticas também poderiam ser considerados. Pode escolher mais do que uma opção."
List of options with checkboxes:
 - Índice de assuntos
 - Nuvem de assuntos
 - Tesouro
 - Ontologia

The browser's taskbar at the bottom shows various application icons and the system tray with the date 23-11-2013 and time 23:10.

Figura 36 - Página do questionário posterior 1 - SinBAD, ou questionário posterior ao teste de usabilidade 1, no *site* QUETTA (parte 1)

This screenshot is identical to Figure 36, showing the same Google Forms interface. In this version, the dropdown menu for question 1 is closed, and the selected option is "Sim". The dropdown menu for question 2 is also closed, with "Sim" selected. The rest of the form, including the description for question 3 and its list of options, remains the same.

Figura 37 - Página do questionário posterior 1 - SinBAD, ou questionário posterior ao teste de usabilidade 1, no *site* QUETTA (parte 2)

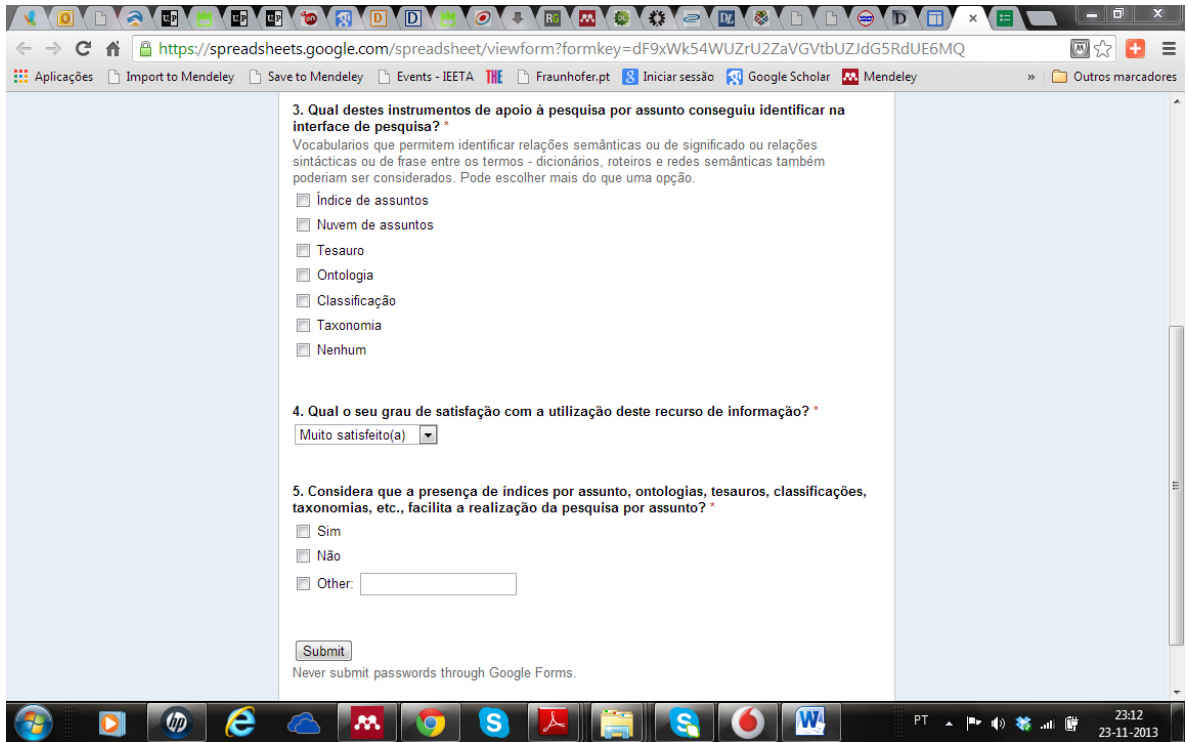


Figura 38 - Página do questionário posterior 1 - SinBAD, ou questionário posterior ao teste de usabilidade 1, no site QUETTA (parte 3)

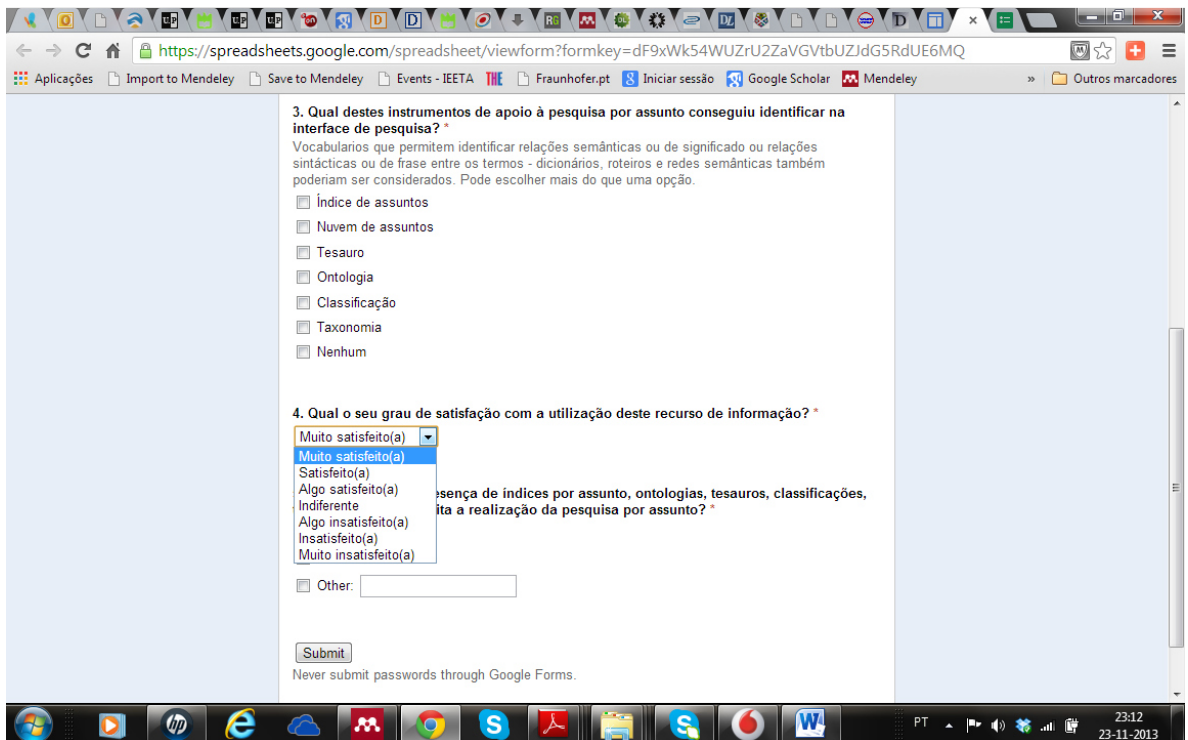


Figura 39 - Página do questionário posterior 1 - SinBAD, ou questionário posterior ao teste de usabilidade 1, no site QUETTA (parte 4)

9. Entretanto, enquanto decorrem os passos 1 a 8, acima enunciados, o observador²¹³ vai tirando notas e cronometrando o comportamento do participante concretizando, desta forma, o *think-aloud 1*;
10. Terminado o preenchimento do questionário posterior 1, o facilitador procede a nova calibragem – **calibragem teste 2** - com o *eyetracker*;
11. Após a calibragem, dá-se início ao teste 2 com as **instruções para o teste 2** na IU do *eyetracker* como se pode verificar na figura 40;
12. O **teste 2** tem início após a leitura das instruções com a ligação ao *site* do repositório digital a ser avaliado: o Repositório Aberto da Universidade do Porto, como se pode ver na figura 41;

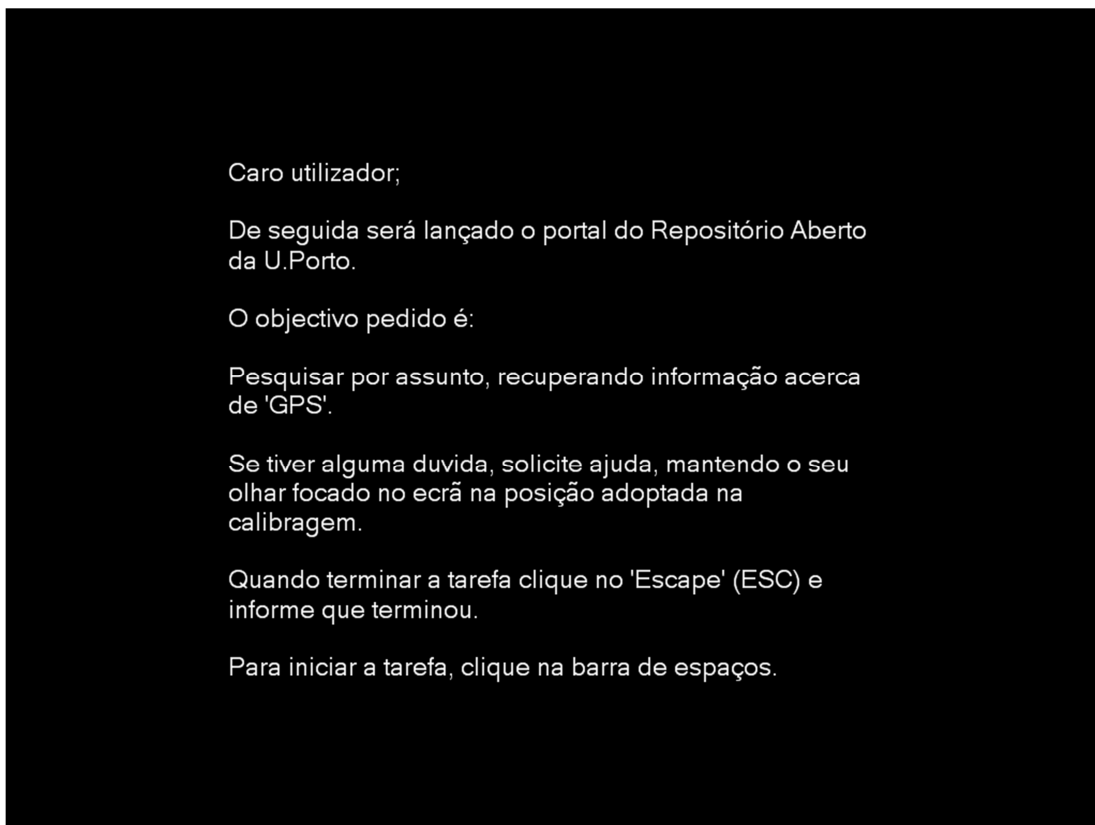


Figura 40 - Instruções para o teste 2 na IU do *eyetracker*

²¹³ O observador desempenhava, igualmente, o papel de facilitador.



Figura 41 – Repositório Aberto da Universidade do Porto (teste 2)

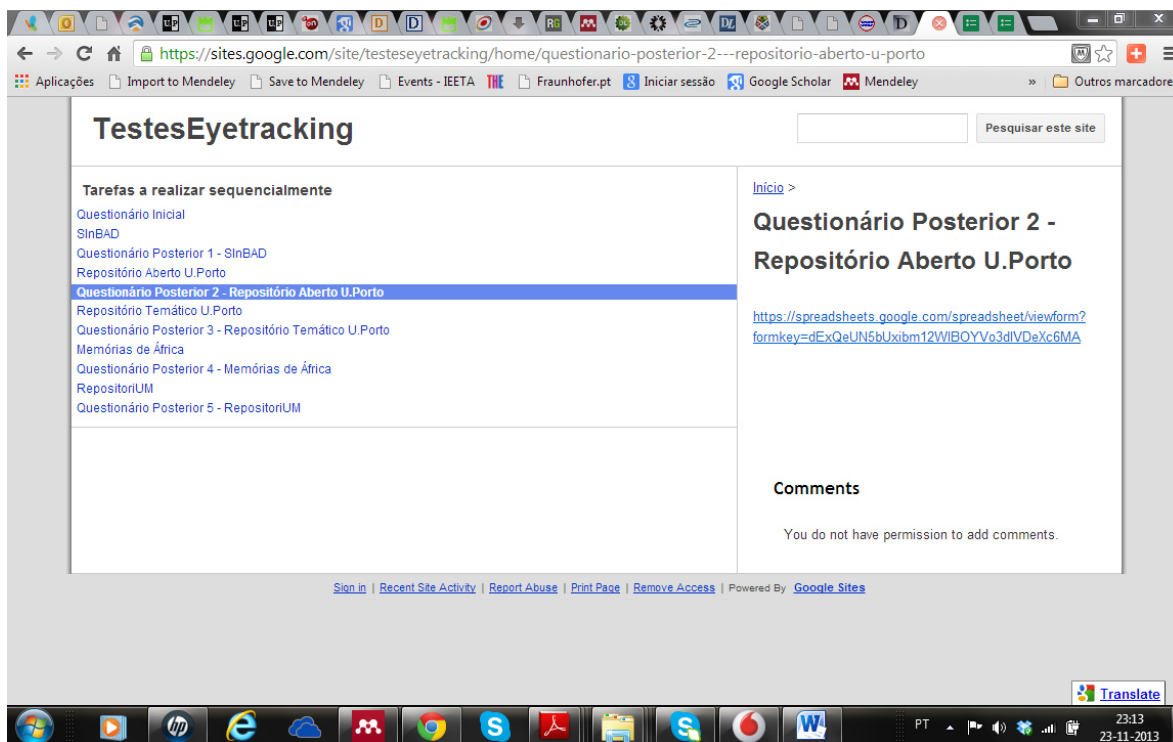


Figura 42 - Página de ligação para o “questionário posterior 2 – Repositório Aberto U.Porto” no site QUETTA

13. Após o teste 2, o participante clica na ligação seguinte na página inicial do *site* QUETTA (cf. figura 42) que o conduz ao **questionário posterior 2** que deverá ser preenchido no final do teste 2 e que consta nas figuras 36, 37, 38 e 39;
14. Entretanto, o observador vai tirando notas e cronometrando o comportamento do participante concretizando, desta forma, o ***think-aloud 2***;
15. Terminado o preenchimento do questionário posterior 2, o facilitador procede a nova calibragem – **calibragem teste 3** - com o *eyetracker*;
16. Após a calibragem, dá-se início ao teste 3 com as **instruções para o teste 3** na IU do *eyetracker* como se pode verificar na figura 43;

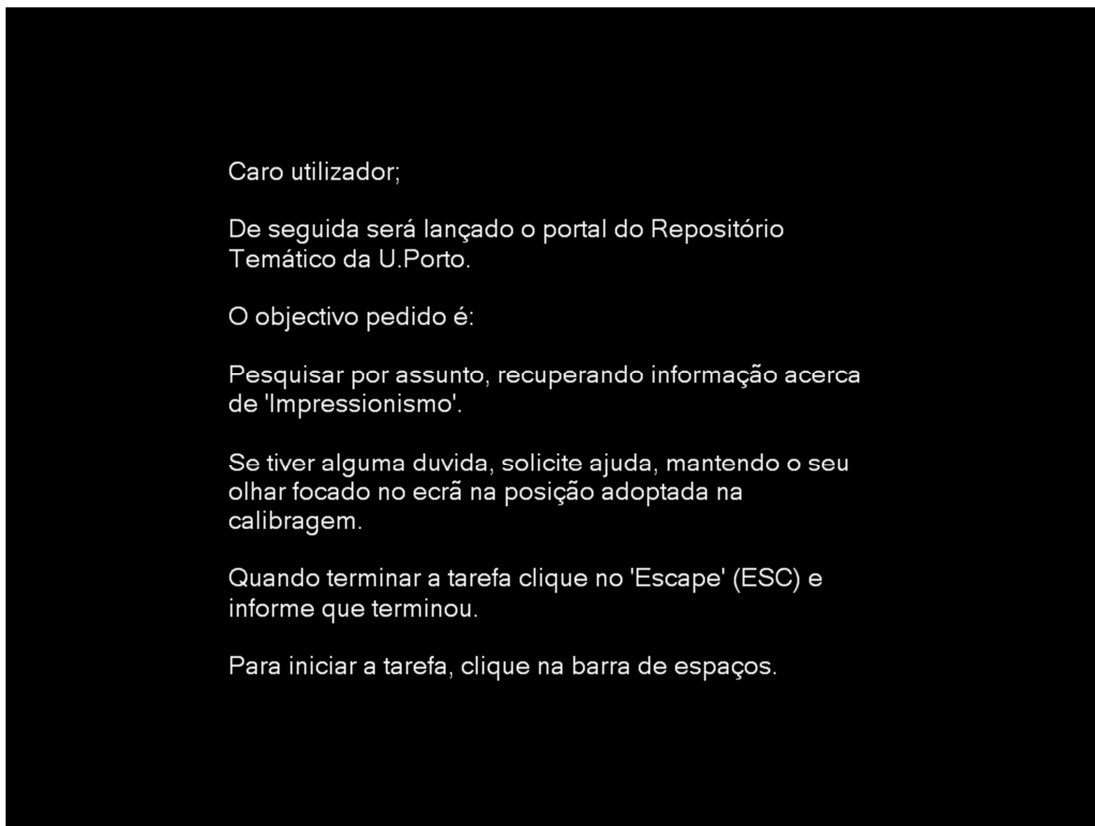


Figura 43 - Instruções para o teste 3 na IU do *eyetracker*



Figura 44 - Repositório Temático da Universidade do Porto (teste 3)



Figura 45 - Página de ligação para o “questionário posterior 3 – Repositório Temático U.Porto” no site QUETTA

17. O **teste 3** tem início após a leitura das instruções com a ligação ao *site* do repositório digital a ser avaliado: o Repositório Temático da Universidade do Porto, como se pode ver na figura 44;
18. Após o teste 3, o participante clica na ligação seguinte na página inicial do *site* QUETTA (cf. figura 45) que o conduz ao **questionário posterior 3** que deverá ser preenchido no final do teste 3 e que consta nas figuras 36, 37, 38 e 39;
19. Enquanto isso, o observador vai tirando notas e cronometrando o comportamento do participante concretizando, desta forma, o **think-aloud 3**;
20. Terminado o preenchimento do questionário posterior 3, o facilitador procede a nova calibragem – **calibragem teste 4** - com o *eyetracker*;
21. Após a calibragem, dá-se início ao teste 4 com as **instruções para o teste 4** na IU do *eyetracker* como se pode verificar na figura 46;

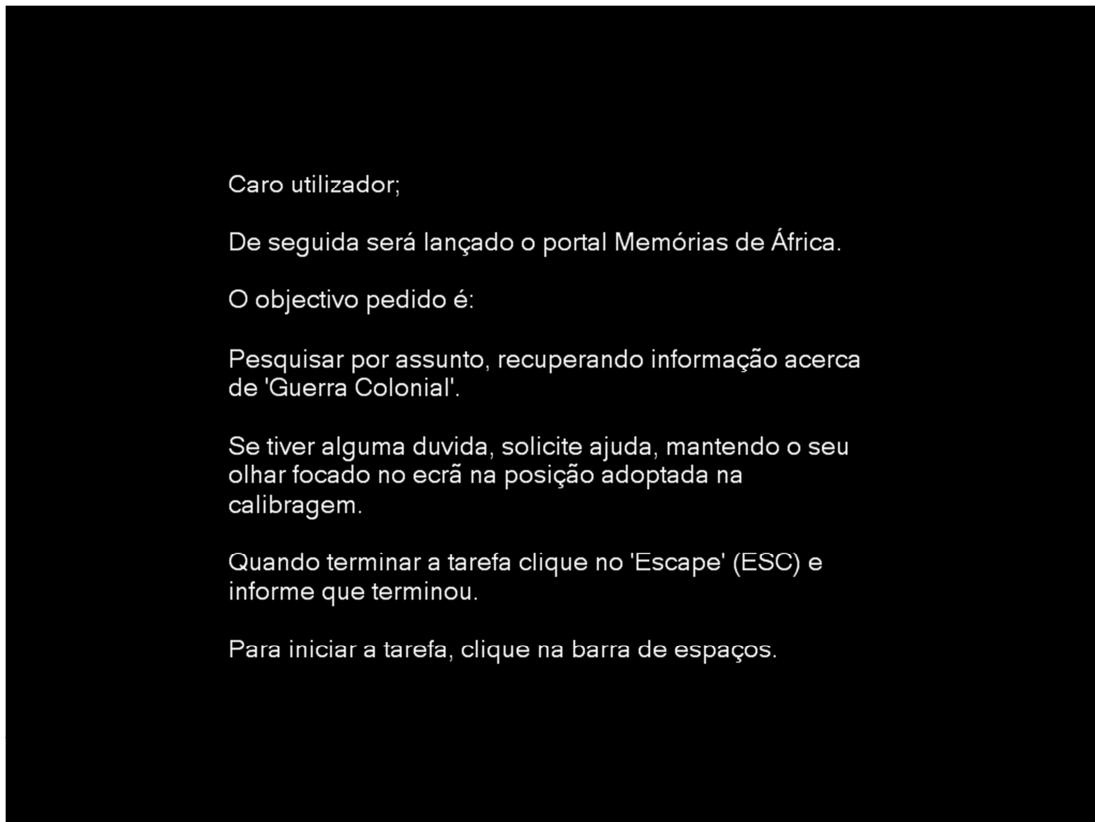


Figura 46 - Instruções para o teste 4 na IU do *eyetracker*



Figura 47 – Memórias de África²¹⁴, da Fundação Portugal-África, mantida pela Universidade de Aveiro e pelo Centro de Estudos sobre África e do Desenvolvimento (teste 4)

22. O teste 4 tem início após a leitura das instruções com a ligação ao *site* do repositório digital a ser avaliado: Memórias de África da Fundação Portugal-África, mantida pela Universidade de Aveiro e pelo Centro de Estudos sobre África e do Desenvolvimento Universidade de Aveiro, como se pode ver na figura 47;
23. Após o teste 4, o participante clica na ligação seguinte na página inicial do *site* QUETTA (cf. figura 48) que o conduz ao **questionário posterior 4** que deverá ser preenchido no final do teste 4 e que consta nas figuras 36, 37, 38 e 39;
24. Entretanto, o observador vai tirando notas e cronometrando o comportamento do participante concretizando, desta forma, o **think-aloud 4**;
25. Terminado o preenchimento do questionário posterior 4, o facilitador procede a nova calibragem – **calibragem teste 5** - com o *eyetracker*;

²¹⁴ Fundada em 1997, mudou de nome após a data de realização deste estudo que decorreu no início de Maio de 2011. O nome atual é “Memórias de África e do Oriente”.

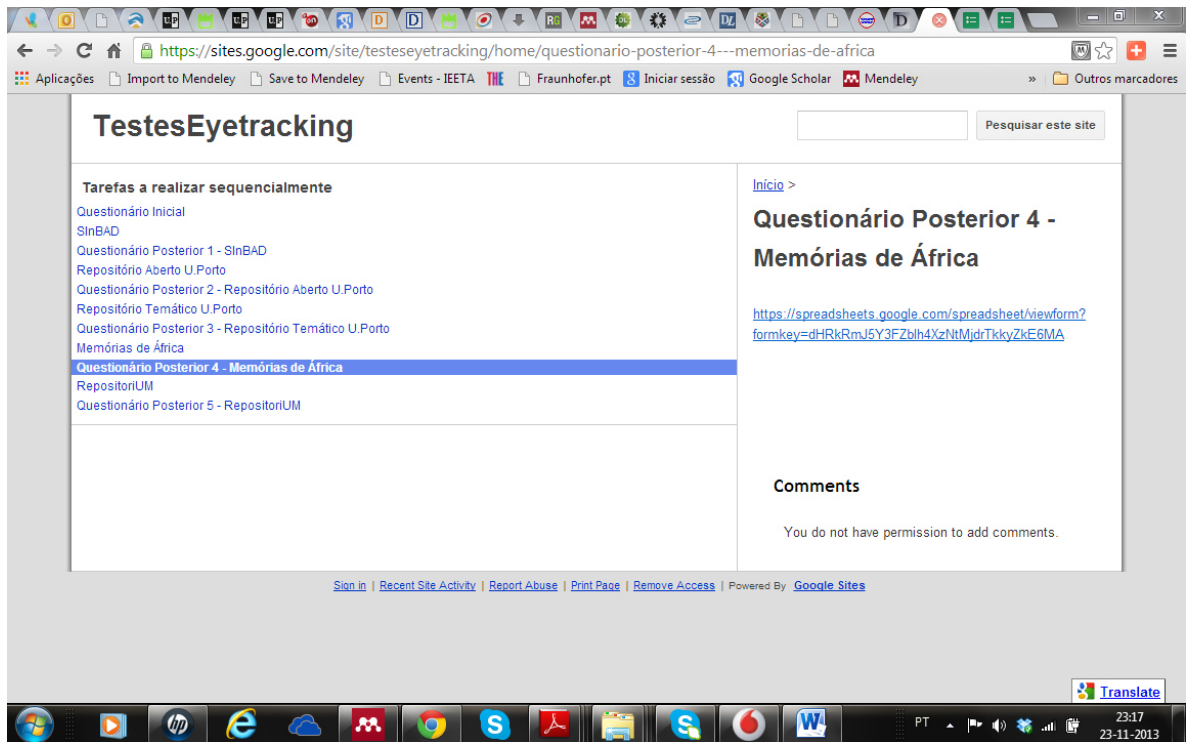


Figura 48 - Página de ligação para o “questionário posterior 4 – Memórias de África” no *site* QUETTA

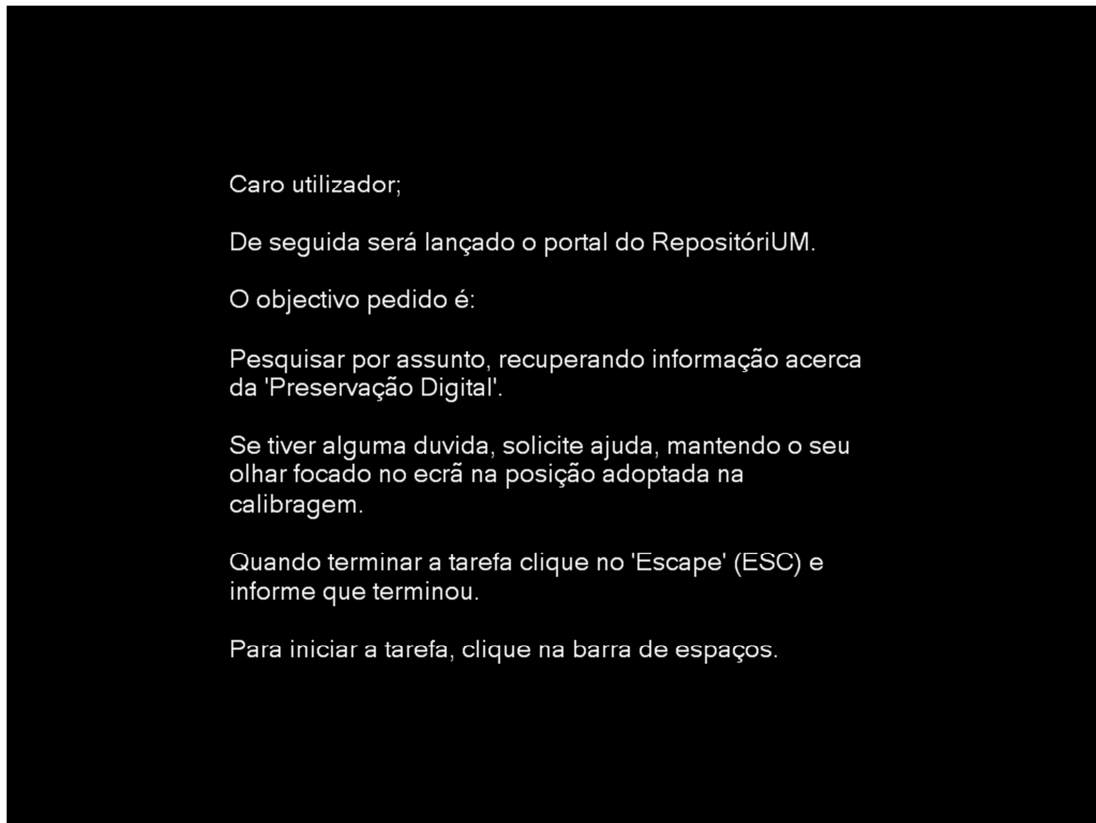


Figura 49 - Instruções para o teste 5 na IU do *eyetracker*



Figura 50 – RepositórioUM da Universidade do Minho (teste 5)

26. Após a calibragem, dá-se início ao teste 5 com as **instruções para o teste 5** na IU do *eyetracker* como se pode verificar na figura 49;
27. O **teste 5** tem início após a leitura das instruções com a ligação ao *site* do repositório digital a ser avaliado: RepositórioUM da Universidade do Minho, como se pode ver na figura 50;
28. Após o teste 5, o participante clica na ligação seguinte na página inicial do *site* QUETTA (cf. figura 51) que o conduz ao **questionário posterior 5** que deverá ser preenchido no final do teste 5 e que consta nas figuras 36, 37, 38 e 39;
29. Durante esses passos, o observador vai tirando notas e cronometrando o comportamento do participante concretizando, desta forma, o **think-aloud 5**;
30. Chega-se ao final do estudo, tendo completado a bateria de testes e os vinte e oito (28) passos do sistema QUETTA, tendo os dados ficado armazenados para posterior tratamento.

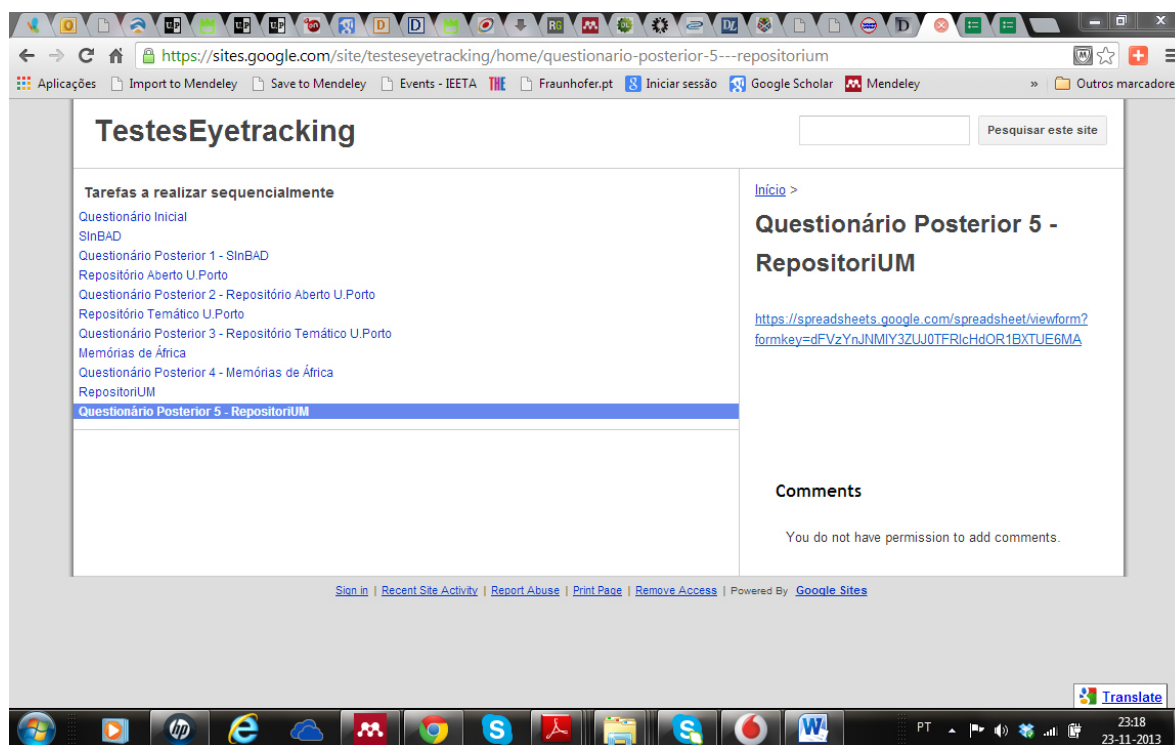


Figura 51 - Página de ligação para o “questionário posterior 5 – RepositoriUM” no *site* QUETTA

3.3. A recolha de dados gerada pelo sistema QUETTA

O estudo do comportamento do utilizador relatado nesta tese decorreu em Abril e Maio de 2011. A preparação do estudo decorreu, ainda, no mês de Abril de 2011 com o recrutamento por *email* dos voluntários participantes no estudo e o planeamento da montagem do mesmo. A montagem decorreu entre a última semana de Abril e os primeiros dias de Maio de 2011 – o *site* QUETTA ficou concluído a 3 de Maio de 2011. O estudo decorreu de 3 a 6 de Maio de 2011²¹⁵, tendo o dia 3 de Maio sido dedicado a testes para verificar se o cenário do estudo estava bem

²¹⁵ Os quatro (4) dias em que, efetivamente, o *eyetracker* e a sala de laboratório foram cedidos a este projeto de investigação. Estava prevista a cedência do equipamento e da sala, igualmente, no dia 29 de Abril de 2011, mas, essa data acabou por ser alocada a outras equipas de investigação.

montado, se o *eyetracker* e o respetivo *software* funcionavam nas melhores condições, e se o sistema QUETTA estava bem planeado e se os dados recolhidos ficavam bem armazenados. O estudo, propriamente dito, teve início a 4 de Maio, inclusive, e decorreu até 6 de Maio, inclusive, com a agenda repleta de sessões nesses 3 dias.

Foram recrutados vinte e cinco (25) voluntários por *email*, embora só vinte e três (23) tenham tido agendamento, pois, dois (2) deles só tinham disponibilidade para a semana seguinte, semana em que o equipamento já teria regressado às instalações da Portugal Telecom em Lisboa. Dos vinte e três (23) voluntários com disponibilidade para participarem no estudo nos dias em que o equipamento estava disponível, dois (2) faltaram (um com aviso por *email*), pelo que o número de participantes no estudo se fixou em vinte e um (21) voluntários que, realmente, participaram no estudo do comportamento do utilizador com o sistema QUETTA.

Os testes de usabilidade com *eyetracking* foram realizados num *eyetracker* da marca Tobii – modelo T120 Eye Tracker - cedido pela Portugal Telecom SGPS SA (PT) ao Centro de Estudos das Tecnologias e Ciências da Comunicação (CETAC.MEDIA) que, por sua vez, o cedeu a este projeto de investigação por solicitação do Instituto de Engenharia Eletrónica e Telemática de Aveiro (IEETA). Estas duas unidades de investigação (o CETAC.MEDIA das Universidades do Porto e de Aveiro e o IEETA da Universidade de Aveiro) e aquela empresa (PT) forneceram os meios necessários para o avanço do projeto e para a realização, atempada (ainda que com mais de um semestre de atraso relativamente ao inicialmente previsto), do trabalho laboratorial previsto no plano de doutoramento.

O trabalho foi realizado nas instalações do CETAC.MEDIA no Departamento de Comunicação e Arte (DeCA) da Universidade de Aveiro. Nas figuras 52, 53, 54 e 55 que se encontram nas páginas que se seguem pode ver-se o *eyetracker* que foi utilizado nos testes de usabilidade com *eyetracking*.

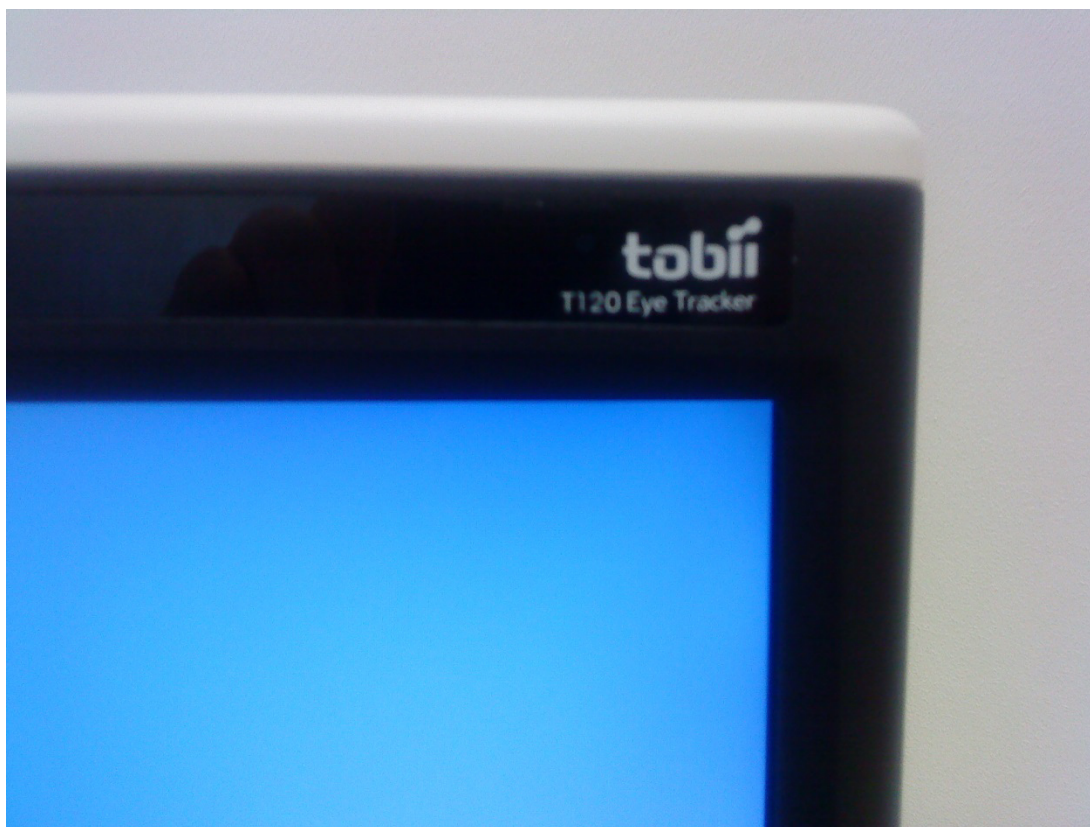


Figura 52 – Eyetracker da marca Tobii (modelo T120 Eye Tracker) – pormenor do canto superior direito – foto de Anabela Serrano

O equipamento para os testes de usabilidade com *eyetracking* foi disposto na sala do laboratório conforme consta na foto identificada como sendo a figura 56. Foram feitos reajustes de posicionamento da cadeira, do *eyetracker*, do rato e do teclado antes de cada sessão, medindo as distâncias de forma a facilitar a calibragem dos olhos com o *eyetracker*, a assegurar o conforto do participante em termos de ergonomia e de visualização, e a permitir a observação do participante pelo observador. A disposição do mobiliário na sala do laboratório foi alterada por causa do protocolo *think-aloud*. Assim, foi colocada uma mesa e uma cadeira à direita da mesa onde estava instalado o *eyetracker*, e atrás da cadeira do participante (em 3ª linha), a uma distância suficiente, e com ângulo de visualização correto, para permitir uma observação exata do comportamento daquele sem que a presença do observador se fizesse sentir, evitando assim qualquer perturbação do comportamento do participante no estudo. Foram feitos reajustes de posicionamento da mesa e da cadeira antes do início de cada bloco de sessões (no início do dia após a passagem da

equipa de limpeza, e, por vezes, no início da tarde se tivesse havido alteração da disposição durante a ausência da observadora no período de almoço).



Figura 53 - Eyetracker da marca Tobii (modelo T120 Eye Tracker) – vista lateral – foto de Anabela Serrano

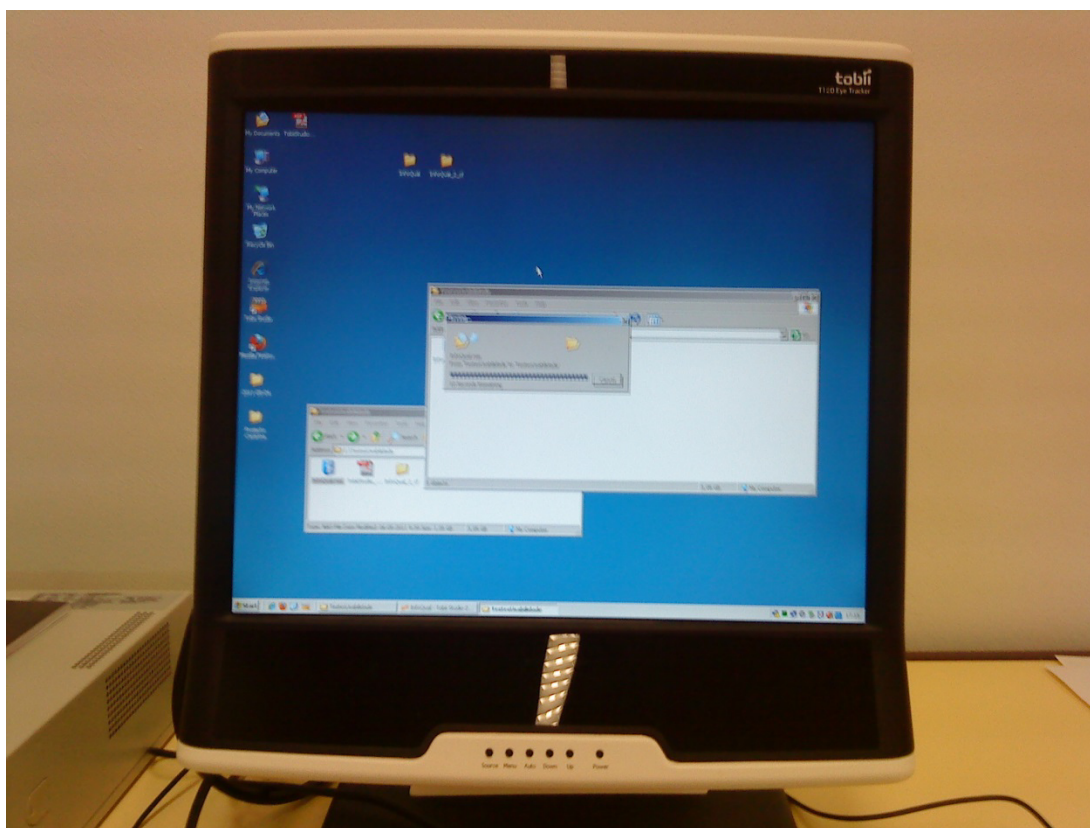


Figura 54 - Eyetracker da marca Tobii (modelo T120 Eye Tracker) – vista frontal 1 – foto de Anabela Serrano

Segundo Nielsen (2000), o número de participantes em testes de usabilidade com *eyetracking* que permite obter resultados fiáveis é de cinco (5) participantes ou utilizadores²¹⁶. Neste estudo colaboraram vinte e um (21) participantes que se podem subdividir em várias categorias, de acordo com o seu contexto académico e/ou socioprofissional, havendo cinco (5) a seis (6) participantes em cada uma delas:

- 5 docentes e/ou investigadores;
- 5 profissionais de informação;
- 6 alunos de doutoramento;
- 5 alunos de mestrado/pós-graduação/mestrado integrado.

²¹⁶ <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
<http://www.nngroup.com/articles/how-many-test-users/>

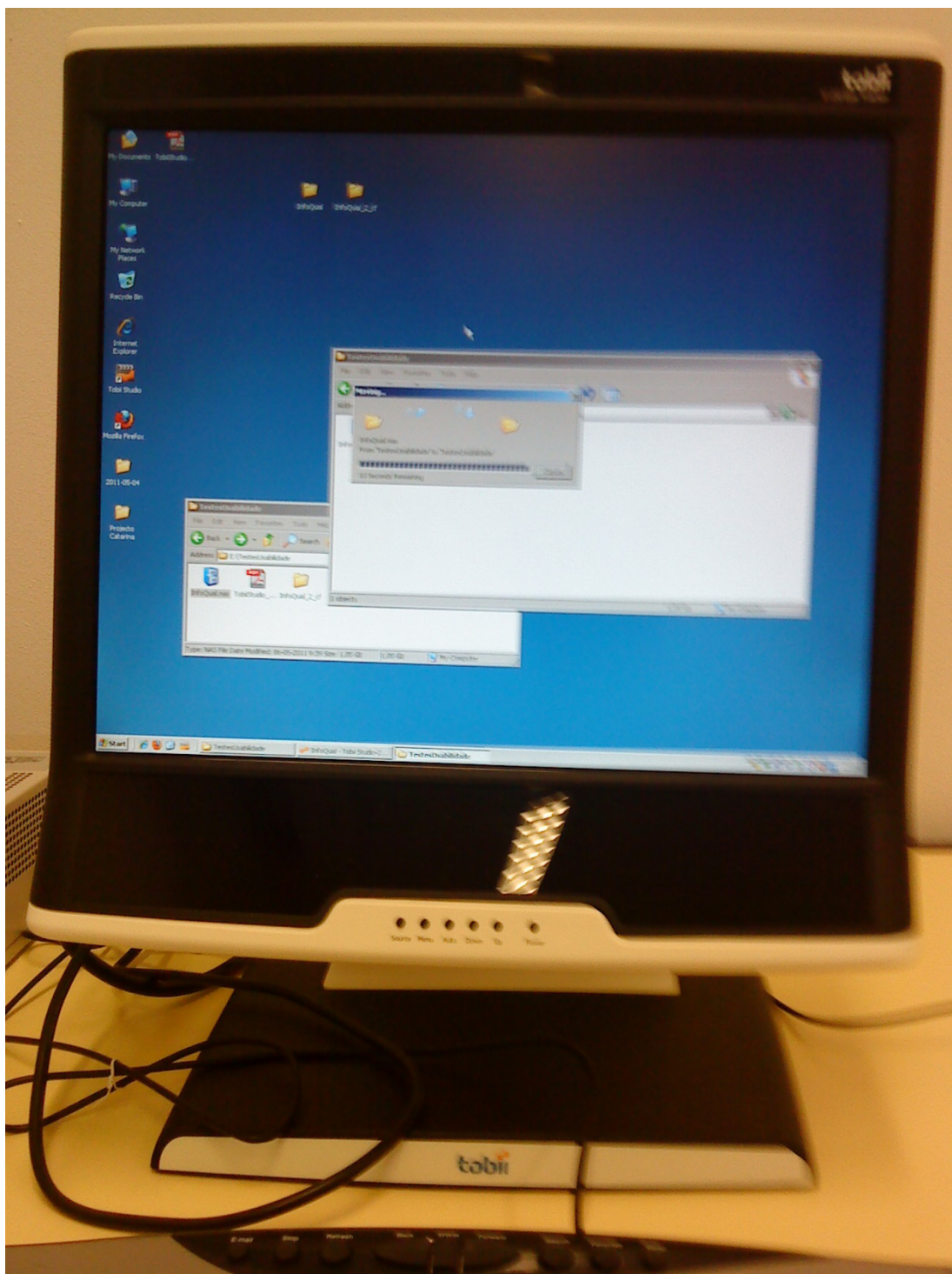


Figura 55 - Eyetracker da marca Tobii (modelo T120 Eye Tracker) – vista frontal 2 – foto de Anabela Serrano



Figura 56 – Disposição dos elementos do equipamento e posição da cadeira do participante frente ao *Eyetracker* da marca Tobii (modelo T120 Eye Tracker) – foto de Anabela Serrano

Quer se analise o total de participantes, quer se analise o número de participantes por categoria, o número de voluntários que participaram no estudo ultrapassou o número estimado como sendo fiável que são cinco (5) utilizadores e o número considerado necessário para identificar todos os problemas de usabilidade que são quinze (15) utilizadores.

Em categorias, segundo Nielsen, basta testar três (3) a quatro (4) utilizadores por cada categoria. Ora, neste estudo, houve cinco (5) participantes na maioria das categorias e seis (6) numa delas. Apenas, não se chegou ao número necessário de participantes que permitiria usar os *heatmaps* como fonte primária de dados tal como os *gaze plots* foram utilizados, que seriam trinta (30) ou mais utilizadores para obter 15% de dados aleatórios ou menos. Com vinte e um (21) utilizadores ter-se-á um pouco menos de 30% de dados aleatórios, ou seja, pouco mais de 70% de

legibilidade nos *heatmaps* ou mapas de calor²¹⁷. Por esse motivo, os *heatmaps* são utilizados, apenas, a título ilustrativo neste estudo. O que não é importante, dado que os *gaze plots* ou mapas de fixações são bem mais elucidativos quanto às fixações e aos percursos seguidos pelos utilizadores na IU.

O sistema QUETTA pode ser sintetizado em números, tal como se pode ver abaixo:

- ✚ 3 técnicas de investigação quantitativas e qualitativas que se incluem em 3 métodos de investigação;
- ✚ 28 componentes ou passos, em que 11 são executados pelo participante, e 17 pelo facilitador/observador;
- ✚ 6 questionários a preencher pelo participante num total de 31 questões;
- ✚ 5 testes de usabilidade com *eyetracking* a serem realizados pelo participante;
- ✚ 5 *think-aloud* levados a cabo pelo observador;
- ✚ 5 calibrações efetuadas pelo facilitador/observador;
- ✚ 6 conjuntos de instruções dadas pelo facilitador;
- ✚ 646 respostas aos questionários recolhidas;
- ✚ 24,6 GB de dados recolhidos com o *software* da Tobii;
- ✚ 100 a 200 fixações por *webpage* por participante;
- ✚ 392 registos de *think-aloud* efetuados.

3.3.1. Os questionários

Os questionários são compostos por seis (6) questões no caso do questionário prévio aos testes de usabilidade com *eyetracking*, e cinco (5) questões no caso dos cinco (5) questionários posteriores a cada um dos cinco (5) testes de usabilidade com *eyetracking*. As questões são do tipo fechado e aberto, embora predominem as questões fechadas. O objetivo do questionário prévio é o de caracterizar os participantes no estudo. O objetivo dos questionários posteriores é o de recolher as primeiras impressões do participante logo após a realização do teste de forma a poder

²¹⁷ Cf. Pernice, Nielsen (2009), p. 43-51.

cruzar as respostas com os registos de vídeo e de imagem do *eyetracker* e com os registos de observação do *think-aloud*. O texto completo dos seis (6) questionários com a indicação de qual o repositório digital em análise nos questionários posteriores aos testes de usabilidade com *eyetracking* consta do anexo 6. As figuras 28, 29, 30, 31 e 32 reproduzem o aspeto do questionário prévio aos testes de usabilidade com *eyetracking* no site QUETTA.

Nas linhas abaixo, reproduz-se o texto do questionário prévio.

Questionário prévio

1. Qual é a sua atividade na Universidade de Aveiro? (Pode escolher mais do que uma opção)

Docente/Investigador
Profissional de Informação
Aluno Doutoramento
Aluno Mestrado/ Pós-Graduação
Aluno Licenciatura/Mestrado Integrado

2. Qual é o seu grupo etário?

18-29 anos
30-39 anos
40-49 anos
50-59 anos
60-69 anos
70-79 anos
+ de 80 anos

3. Qual é o seu nível de habilitações académicas?

Ensino secundário (12º ano)
Ensino Pós-Secundário (CET)
Bacharelato
Licenciatura
Mestrado
Doutoramento
Outro. Qual? _____

4. Costuma fazer pesquisa de informação por assunto em plataformas digitais? (Se respondeu “não” o questionário prévio termina aqui)

Sim
Não

5. Qual a frequência com que acede a esses recursos de informação?

Frequentemente
Ocasionalmente

6. A pesquisa por assunto é para si mais usual do que a pesquisa por autor e título?

Sim
Não

As figuras 36, 37, 38 e 39 reproduzem o aspeto do questionário posterior ao primeiro teste de usabilidade com *eyetracking* no *site* QUETTA.

Nas linhas abaixo reproduz-se o texto do questionário posterior.

Questionário posterior ao Teste de Usabilidade 1

ID: User__

Recurso de Informação - SInBAD:

<http://biblioteca.sinbad.ua.pt/Pesquisa/?sub=4>

1. Considera que foi fácil aceder à informação pretendida?

Sim
Não

2. Encontrou auxílio para a pesquisa por assunto na interface de pesquisa?

Sim
Não

3. Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa? (Pode escolher mais do que uma opção)

Índice de assuntos
Classificação

Nuvem de assuntos
Ontologia
Taxonomia
Tesauro
Nenhum

4. Qual o seu grau de satisfação com a utilização deste recurso de informação?

Muito satisfeito(a)
Satisfeito(a)
Algo satisfeito(a)
Indiferente
Algo insatisfeito(a)
Insatisfeito (a)
Muito insatisfeito(a)

5. Considera que a presença de índices por assunto, ontologias, tesauros, classificações, taxonomias, etc., facilita a realização da pesquisa por assunto?

Sim
Não
Outro _____

Das seiscentas e cinquenta e uma (651) respostas passíveis de serem recolhidas *online* só foram submetidas seiscentas e quarenta e seis (646) respostas, havendo um participante que falhou a resposta a cinco (5) questões.

3.3.2. Os mapas de fixações e os *heatmaps*

Os testes de usabilidade com *eyetracking* deram origem a 24,6 GB de dados que podem ser visualizados em mapas de fixações ou *gaze plots* (os que parecem ser mais interessantes e úteis para este estudo), em mapas de calor ou *heatmaps* (os que atraem mais as pessoas sob o ponto de vista visual, de acordo com o testemunho de especialistas²¹⁸), e em vídeos. Na figura 57, pode ver-se um mapa de fixações com a sobreposição das fixações de vários participantes, as quais podem ser distinguidas pelas diferentes cores atribuídas a cada ID de utilizador. Boa parte dos mapas de

²¹⁸ Cf. Pernice, Nielsen (2009).

fixações de cada participante têm mais de cem (100) fixações e, nalguns casos, mais de duzentas (200) fixações numa única *webpage*. O número de *webpages* percorridas ao longo de um teste pode ser superior a vinte (20), pelo que, facilmente, se percebe o impressionante volume de dados que foi gerado durante os testes e as inúmeras relações possíveis entre os dados dos vários participantes. Para além disso, as gravações em vídeo de cada sessão feitas pelo *eyetracker* permitem visualizar o movimento do olhar do participante e visualizar os fotogramas ao mais ínfimo pormenor. Os dados recolhidos são muitíssimo ricos, pelo que as possibilidades de tratamento dos mesmos são inúmeras.

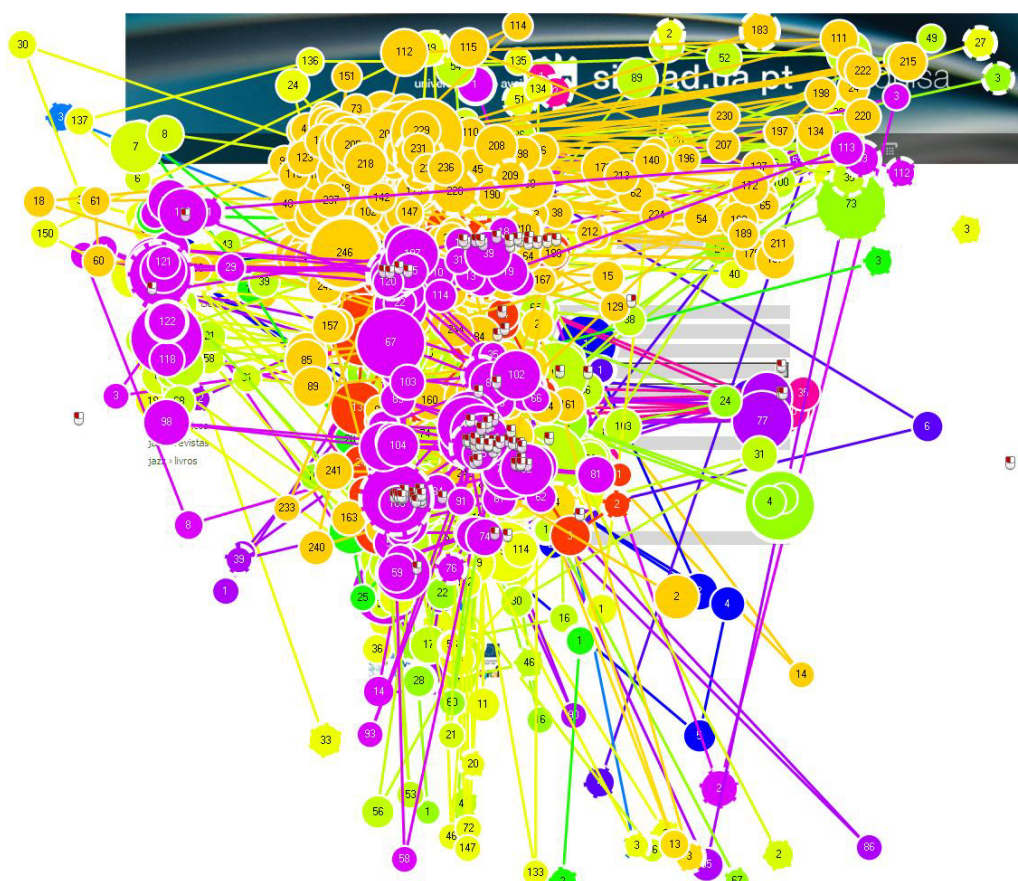


Figura 57 – Exemplo de um mapa de fixações ou *gaze plot* com as fixações de vários participantes sobrepostas e diferenciadas pela atribuição de uma cor diferente a cada utilizador

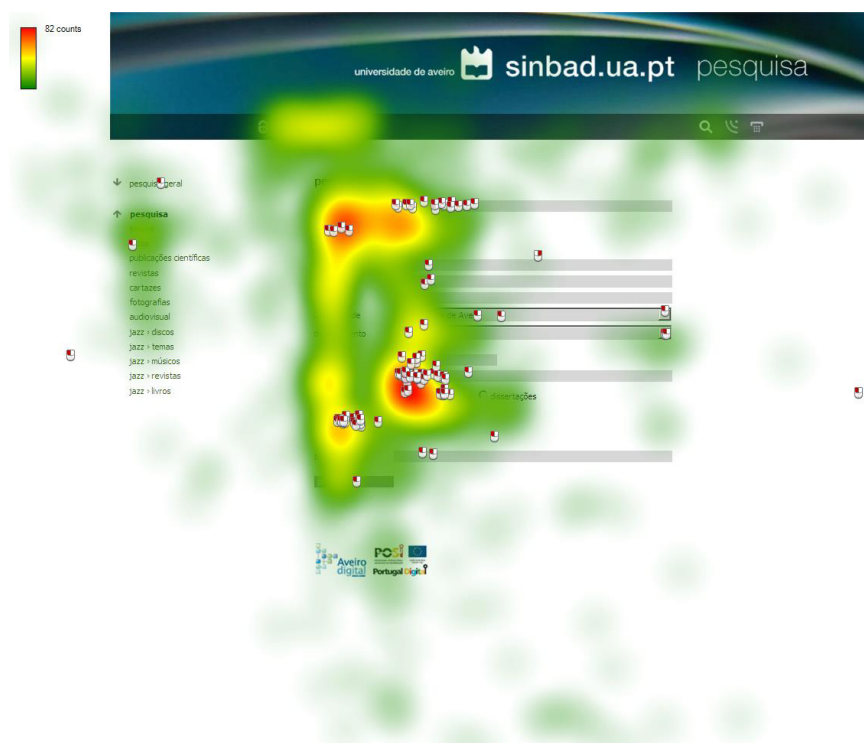


Figura 58 – Exemplo de *heatmap* por contagem de cliques

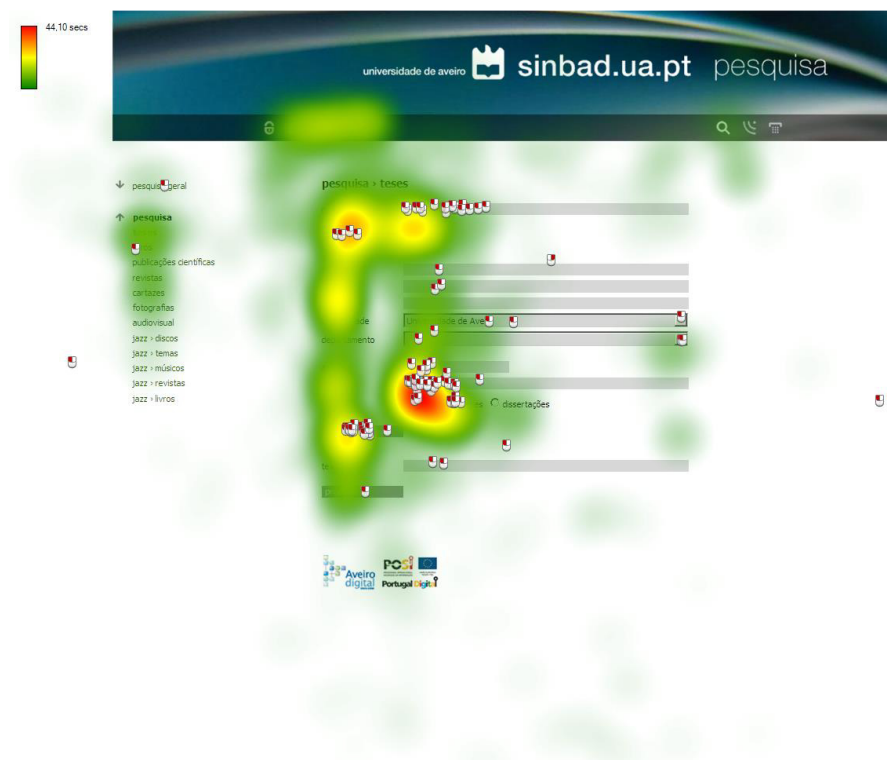


Figura 59 - Exemplo de *heatmap* por duração absoluta

Os participantes receberam instruções escritas acerca das tarefas a executarem durante os testes de usabilidade com *eyetracking*. As cinco instruções redigidas para cada teste solicitavam que os participantes submetessem uma expressão de pesquisa por assunto bastante simples, dado que o objetivo não era avaliar a capacidade do utilizador para a resolução de problemas de recuperação da informação, mas sim avaliar a presença e o papel dos SOC na IU dos SRI dos repositórios digitais selecionados para o estudo, bem como o comportamento do utilizador durante a pesquisa com ou sem o auxílio dos SOC.

Os textos das instruções para os testes 1 a 5 encontram-se reproduzidos, respetivamente, nas figuras 33, 40, 43, 46 e 49.

3.3.3. O *think-aloud*

Os cinco (5) *think-aloud* resultaram em trezentos e noventa e dois (392) registos efetuados pela observadora. Para efetuar os registos foi utilizado um computador portátil e um cronómetro digital.

De forma a minimizar ao máximo a eventual influência da observadora no comportamento do participante, não foi pedido que aquele falasse em voz alta, mas antes foi explicado que deveria atuar da forma mais natural e confortável possível, tal como procederia se estivesse sozinho, pelo que se quisesse falar ou murmurar ou protestar, toda essa expressividade seria útil para o estudo tal como o seu inverso. Ou seja, o importante era tentar abstrair-se da presença da facilitadora/observadora para que a pesquisa solicitada decorresse com a maior naturalidade. O que se tentou com este esclarecimento foi evitar, da melhor forma possível, o condicionamento do comportamento do participante e o aumento do tempo de execução da pesquisa. De igual modo, os mapas de fixações e os *heatmaps* saem mais puros se não se pedir para o utilizador falar ao longo do teste de usabilidade com *eyetracking*.

A cronometragem foi aplicada desde o início do preenchimento do questionário prévio até à submissão do questionário posterior 5. Mas, como o *think-aloud* foi interrompido sempre que era necessário efetuar as calibrações, dado que a observadora era, também, a facilitadora, no modelo indicam-se cinco (5) *think-aloud*, pois, as anotações eram interrompidas nos períodos dedicados à calibragem, salvo os registos do cronómetro que foram sempre efetuados.

Capítulo 4

Análise dos dados e discussão dos resultados da investigação

4.1. Análise dos dados

Os seis (6) questionários que constam do anexo 6, incluíam trinta e uma (31) questões no seu conjunto e obtiveram seiscentas e quarenta e seis (646) respostas, como já foi mencionado no capítulo anterior.

Neste ponto do capítulo 4, serão analisados os dados recolhidos através do sistema QUETTA, subdivididos pelas três técnicas de investigação utilizadas na composição do mesmo: questionários, testes de usabilidade com *eyetracking*, e *think-aloud*. Assim, em primeiro lugar, analisam-se os dados recolhidos através dos questionários. A ordem de apresentação dos dados seguirá a ordem pela qual foram recolhidos durante o estudo, ou seja, a sequência delineada no sistema QUETTA.

4.1.1. Questionários

Inicia-se a análise pelo questionário prévio, seguindo-se o questionário posterior. Após a análise do questionário prévio, e antes da análise do questionário posterior, apresenta-se o perfil do participante sob o ponto de vista demográfico (ponto 4.1.1.1.) e sob o ponto de vista das competências informacionais (ponto 4.1.1.2.), em resultado de uma síntese da análise dos dados do questionário prévio, o que poderá permitir uma melhor compreensão da análise de dados dos cinco questionários posteriores aos testes de usabilidade com *eyetracking* 1, 2, 3, 4 e 5.

Questionário prévio – caracterização dos participantes²¹⁹ no estudo

O questionário prévio tem por objetivo caracterizar os participantes no estudo realizado com o sistema QUETTA. Os participantes foram caracterizados quanto a:

- atividade na Universidade de Aveiro onde foi efetuado o recrutamento dos voluntários e onde foi realizado o estudo em ambiente laboratorial;
- idade;
- habilitações académicas;
- experiência em pesquisa de informação por assunto em plataformas digitais;
- frequência de acesso à informação em plataformas digitais;
- predominância da pesquisa por assunto ou da pesquisa por autor e título no acesso à informação.

Questão 1 – “Qual é a sua atividade na Universidade de Aveiro?”

O estudo teve um total de vinte e um (21) participantes voluntários, de entre os quais sete (7) têm mais de um perfil: três (3) dos participantes que selecionaram a categoria “docente / investigador” selecionaram, também, a categoria “aluno de doutoramento”

²¹⁹ Por uma questão de uniformização, os inquiridos serão designados de “participantes” e não de “respondentes”, dado que participam num estudo que se baseia no sistema QUETTA, o qual é híbrido quanto a métodos de investigação.

e quatro (4) dos participantes que selecionaram a categoria “profissional de informação” selecionaram, igualmente, a categoria “aluno de doutoramento” (1), “aluno de mestrado / pós-graduação” (2) e “aluno de licenciatura / mestrado integrado” (1).

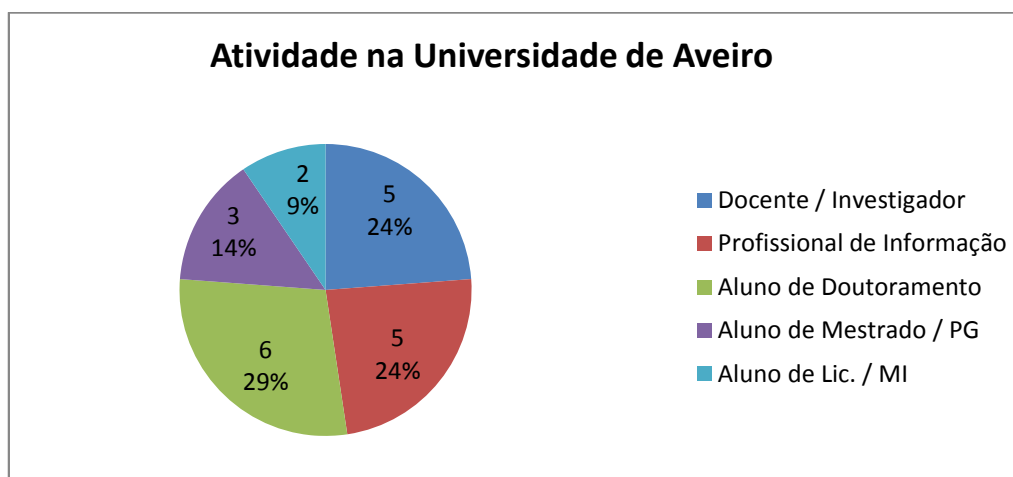


Figura 60 – Gráfico da atividade dos participantes na Universidade de Aveiro

Dado que foi permitida a resposta múltipla nesta questão²²⁰, e que se considerou que a experiência como aluno universitário de 1º, 2º ou 3º ciclo poderia ter influência no comportamento do utilizador, foi decidido optar por uma das opções assinaladas no caso das respostas múltiplas alocando os participantes a uma única categoria²²¹. Assim, dos oito (8) participantes que assinalaram a categoria “docente / investigador”, foram considerados cinco (5) participantes nessa categoria, sendo os restantes três (3) alocados à categoria “aluno de doutoramento”. De entre os nove (9) participantes que selecionaram a categoria “profissional de informação”, cinco (5) participantes foram mantidos nessa categoria, um foi alocado à categoria “aluno de doutoramento” que ficou com um total de seis (6) participantes, dois (2) foram alocados à categoria “aluno de mestrado / pós-graduação” que ficou com um total de

²²⁰ A possibilidade de assinalar mais do que uma opção nesta questão foi intencional, e foi devidamente ponderada antes da realização do estudo, exatamente por se considerar que o comportamento informacional se altera em função da nossa atividade profissional e académica.

²²¹ A questão era de resposta múltipla, pelo que vários participantes assinalaram mais do que uma opção. Os docentes e investigadores e os profissionais de informação que eram, igualmente, alunos de cursos de 1º, 2º e 3º ciclo na Universidade de Aveiro foram alocados às categorias de “aluno de doutoramento”, “aluno de mestrado / pós-graduação” ou de “aluno de licenciatura / mestrado integrado”, consoante o caso.

três (3), e um (1) foi alocado à categoria “aluno de licenciatura / mestrado integrado” que ficou com um total de dois (2).

Como se pode verificar na figura 60, os vinte e um (21) participantes no estudo assinalaram, no seu conjunto, todas as cinco (5) opções de resposta (nalguns casos, duas opções por participante em vez de uma) à questão 1 no questionário prévio, ficando distribuídos da seguinte forma pelas várias categorias:

- 29% dos participantes (6) na categoria “aluno de doutoramento”;
- 24% dos participantes (5) na categoria “docente / investigador”;
- 24% dos participantes (5) na categoria “profissional de informação”;
- 14% dos participantes (3) na categoria “aluno de mestrado / pós-graduação”;
- 9% dos participantes (2) na categoria “aluno de licenciatura / mestrado integrado”.

Questão 2 – “Qual é o seu grupo etário?”

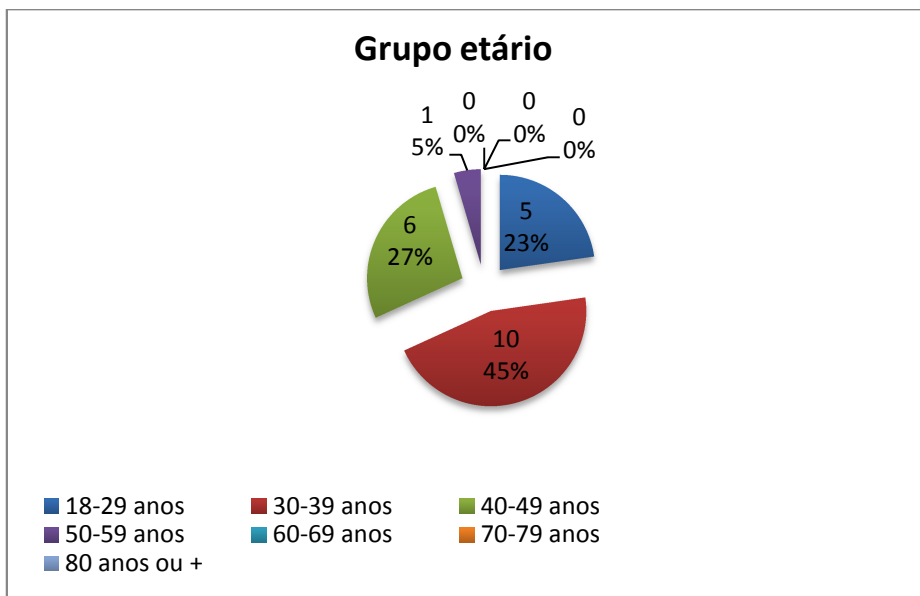


Figura 61 - Gráfico do grupo etário dos participantes

Os vinte e um (21) participantes no estudo assinalaram quatro (4) dos sete (7) grupos etários indicados como opção de resposta à questão 2 no questionário prévio, ficando distribuídos da seguinte forma pelas várias categorias (ver figura 61):

- 45% dos participantes (10) no grupo etário “30-39 anos”;
- 27% dos participantes (6) no grupo etário “40-49 anos”;
- 23% dos participantes (5) no grupo etário “18-29 anos”;
- 5% dos participantes (1) no grupo etário “50-59 anos”;
- 0% dos participantes (0) no grupo etário “60-69 anos”;
- 0% dos participantes (0) no grupo etário “70-79 anos”;
- 0% dos participantes (0) no grupo etário “+ de 80 anos”.

Questão 3 - “Qual é o seu nível de habilitações académicas?”

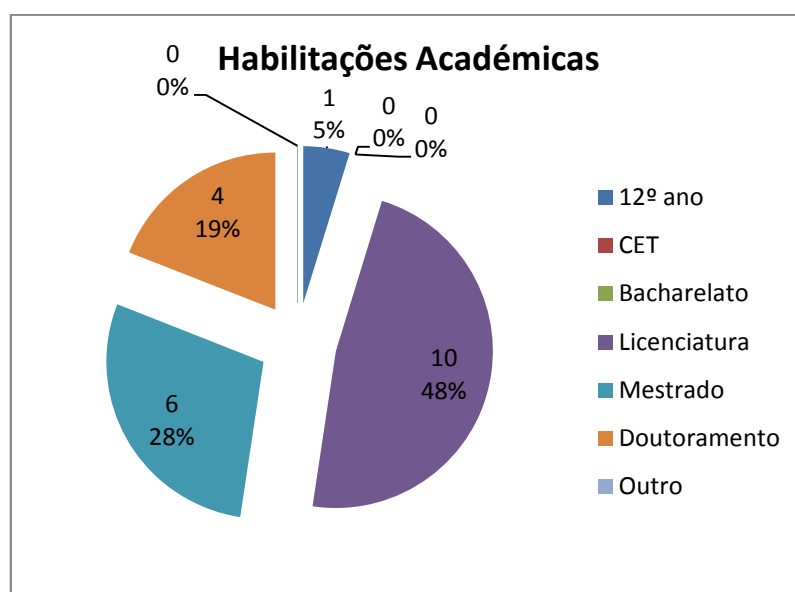


Figura 62 - Gráfico do nível de habilitações académicas dos participantes

Os vinte e um (21) participantes no estudo assinalaram quatro (4) das sete (7) categorias indicadas como opção de resposta à questão 3 no questionário prévio, ficando distribuídos da seguinte forma pelas várias categorias (ver figura 62):

- 48% dos participantes (10) na categoria “Licenciatura”;
- 28% dos participantes (6) na categoria “Mestrado”;
- 19% dos participantes (4) na categoria “Doutoramento”;
- 5% dos participantes (1) na categoria “Ensino Secundário (12º ano)”;
- 0% dos participantes (0) na categoria “CET”;
- 0% dos participantes (0) na categoria “Bacharelato”;
- 0% dos participantes (0) na categoria “Outro”.

Questão 4 - “Costuma fazer pesquisa de informação por assunto em plataformas digitais?”

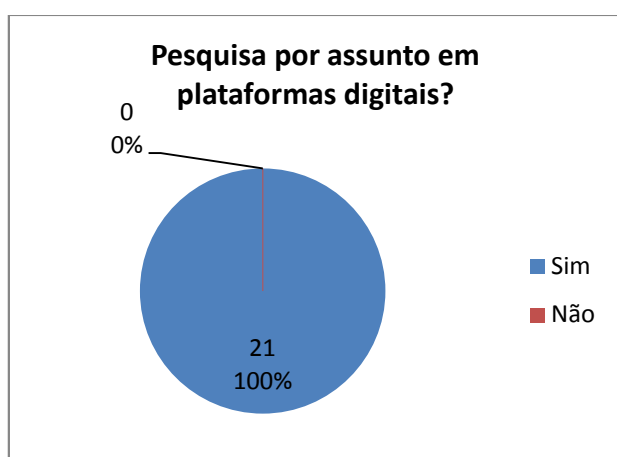


Figura 63 - Gráfico relativo à pesquisa por assunto em plataformas digitais pelos participantes

A resposta a esta questão é unânime conforme se pode verificar na figura 63: todos os vinte e um (21) participantes no estudo assinalaram a opção “Sim”, ou seja, todos costumam pesquisar por assunto nas plataformas digitais.

Questão 5 - “Qual a frequência com que acede a esses recursos de informação?”

A resposta a esta questão é quase unânime como se pode ver na figura 64: vinte (20) dos vinte e um (21) participantes no estudo assinalaram a opção “Frequentemente”, ou seja, quase todos costumam aceder frequentemente às plataformas digitais.

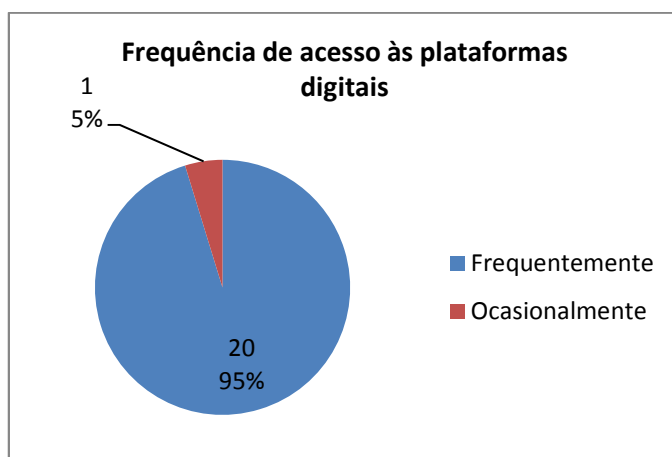


Figura 64 - Gráfico relativo à frequência de acesso às plataformas digitais pelos participantes

Questão 6 - “A pesquisa por assunto é para si mais usual do que a pesquisa por autor e título?”

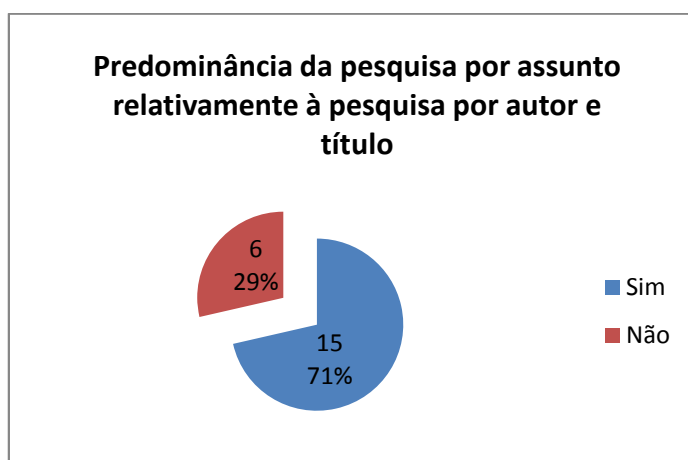


Figura 65 - Gráfico relativo à frequência de pesquisa por assunto comparativamente à pesquisa por autor e título pelos participantes

Quanto à predominância da pesquisa por assunto (ver figuras 65 e 66), dos vinte e um (21) participantes no estudo, quinze (15) assinalaram a opção “Sim” e seis (6) a opção “Não” de resposta à questão 6 no questionário prévio. Ou seja, a pesquisa por assunto predomina, mas, 29% continua a realizar pesquisas por autor/título com maior frequência do que pesquisas por assunto. A única categoria em que só se obtiveram respostas afirmativas foi a dos docentes ou investigadores.

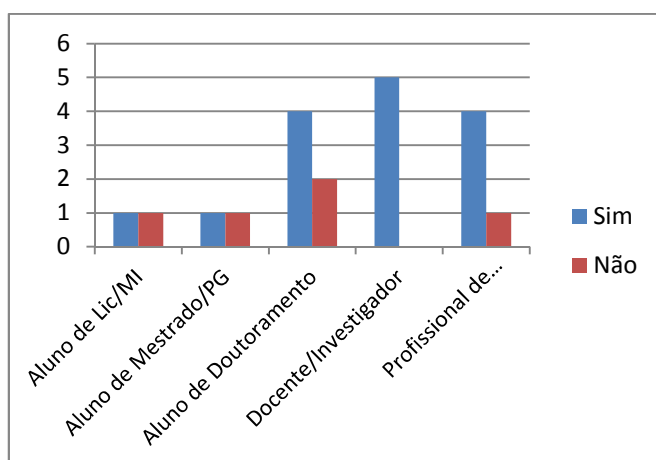


Figura 66 – Gráfico comparativo entre a frequência da pesquisa por assunto e a atividade na Universidade de Aveiro

4.1.1.1. Caracterização demográfica do participante no estudo QUETTA

O perfil do participante²²² no estudo, quanto à sua atividade na Universidade de Aveiro, foi, maioritariamente²²³, de aluno de doutoramento (29%), de docente e investigador (24%) e de profissional de informação (24%), ou seja, o utilizador habituado a pesquisar informação²²⁴ em plataformas digitais quer por exigência da atividade académica e/ou profissional quer pela formação académica e pelas competências informacionais adquiridas.

Quanto à idade, o perfil do participante no estudo situou-se acima dos 18 anos e abaixo dos 60 anos, e foi na sua maioria o de 30 a 39 anos (45% dos voluntários). Os outros dois grupos etários predominantes no desenho do perfil do participante foram o de 40 a 49 anos (27%) e o de 18 a 29 anos (23%). Houve, ainda, um participante

²²² No questionário de caracterização não foi perguntado o género por não ter sido considerado um condicionante do comportamento do utilizador quanto ao objetivo do estudo. Não obstante, regista-se que o perfil do participante foi bastante equilibrado quanto ao género, tendo havido 57% de participantes do sexo feminino (12) e 43% de participantes do sexo masculino (9). Cf. Ghiglione, Matalon (1993), p. 312-315.

²²³ Estas três categorias representam 77% dos participantes.

²²⁴ Poderá dizer-se, mesmo, que em muitos casos dependem da pesquisa da informação e do mundo digital para desenvolverem a sua atividade e até, nalguns casos, na sua vida particular.

com idade entre os 50 e os 59 anos (5%). Os grupos etários acima dos 59 anos não tiveram representação entre os participantes no estudo.

Quanto às habilitações académicas, o perfil do participante foi o de licenciado na maioria dos casos (48% dos voluntários), seguindo-se o de mestre em 28% dos casos, e o de doutor em 19% dos casos. Apenas, 5% dos voluntários (1) tinham por habilitações académicas o ensino secundário (12º ano).

4.1.1.2. Caracterização das competências informacionais do participante no estudo QUETTA

Do ponto de vista informacional, o participante no estudo acede frequentemente (95%) a plataformas digitais, costuma pesquisar por assunto nas mesmas (100%) e faz mais pesquisa por assunto (71%) do que por autor e título (29%).

Questionário posterior

Questão 1 - “Considera que foi fácil aceder à informação pretendida?”

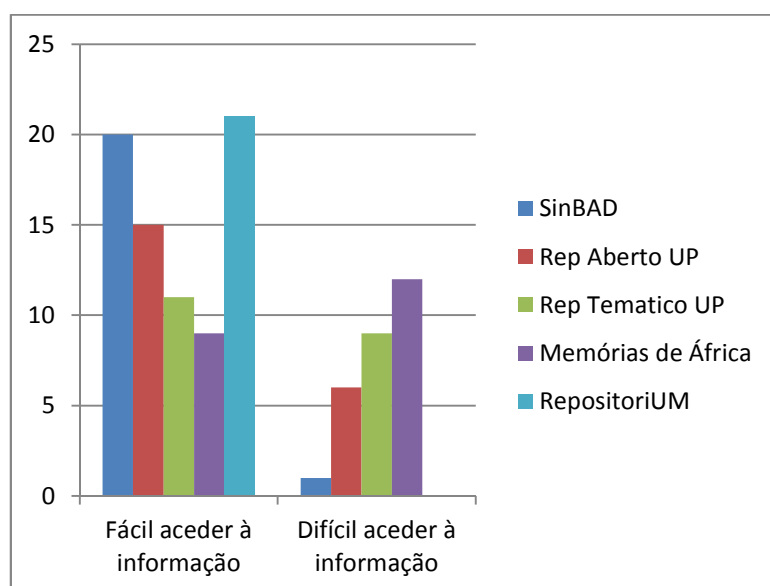


Figura 67 - Gráfico comparativo entre os diferentes repositórios digitais quanto à facilidade de acesso à informação

No que diz respeito à facilidade de acesso à informação, o repositório melhor posicionado foi o RepositóriUM que só obteve respostas positivas (100%). E a SinBAD foi a segunda melhor posicionada com uma quase totalidade de respostas positivas (95%), apenas, uma resposta negativa (5%). De entre os cinco (5) repositórios digitais avaliados todos obtiveram resultados positivos à exceção da “Memórias da África” que obteve mais opiniões negativas (57%) do que positivas (43%). Os dados constam da figura 67.

Questão 2 - “Encontrou auxílio para a pesquisa por assunto na interface de pesquisa?”

Quanto ao auxílio disponível na IU para apoio à pesquisa por assunto, o repositório melhor classificado foi o Repositório Aberto da Universidade do Porto, com uma ligeira diferença positiva - 13 respostas positivas (62%) - relativamente à maioria dos repositórios (3) que obtiveram maior número de respostas positivas do que de respostas negativas – verificar a figura 68.

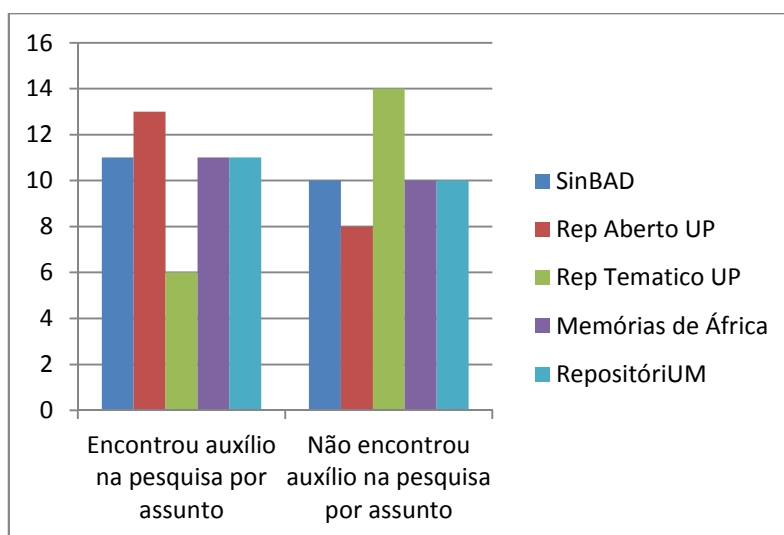


Figura 68 - Gráfico comparativo entre os diferentes repositórios digitais quanto à existência/localização de auxílio à pesquisa por assunto na IU

A SinBAD, a Memórias de África e o RepositóriUM ficaram, *exaequo*, em segundo lugar com onze (11) respostas positivas (52%). O Repositório Temático da

Universidade do Porto²²⁵ foi o único a obter mais respostas negativas do que positivas (30%) que foram, apenas, seis (6).

Questão 3 – “Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa?”

Os participantes identificaram vários SOC nos repositórios digitais, não obstante o facto de nem todos terem sido bem identificados e de alguns nem estarem, mesmo, presentes naquele repositório digital²²⁶. Pela observação feita em simultâneo com o *think-aloud*, tornou-se evidente o desconhecimento de alguns dos conceitos associados à terminologia apresentada no questionário por parte de alguns dos participantes.

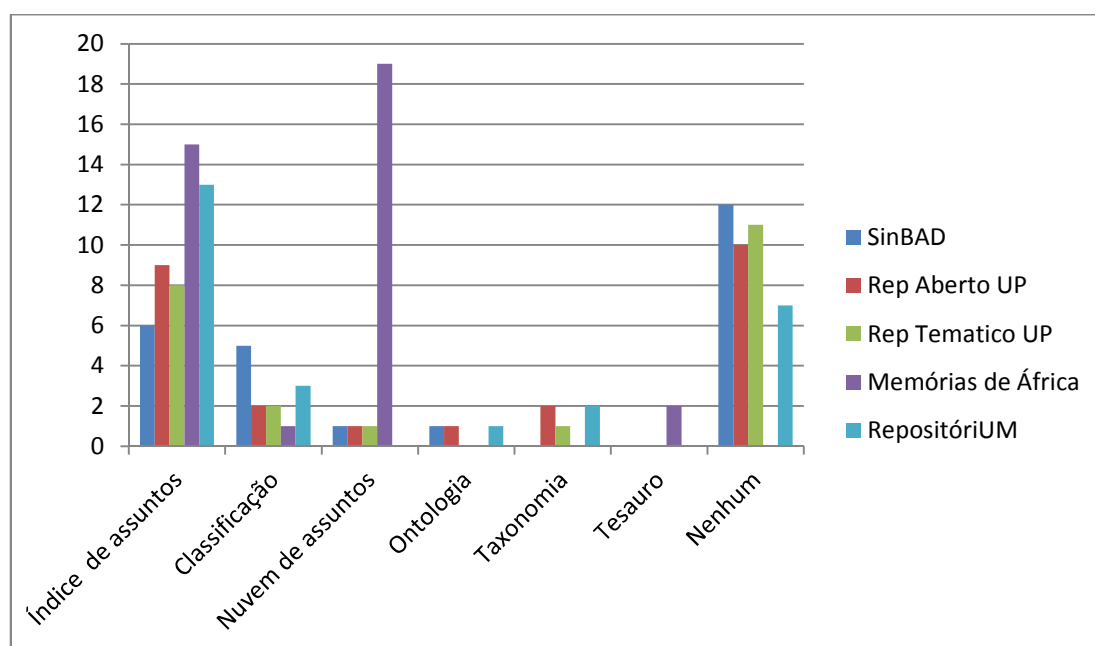


Figura 69 - Gráfico comparativo entre os diferentes repositórios digitais quanto à identificação de SOC na IU

A nuvem de assuntos da Memórias de África foi, nitidamente, o SOC que não causou equívocos e que obteve maior número de identificações com dezanove (19)

²²⁵ O Repositório Temático da Universidade do Porto não obteve respostas por parte de um participante, pelo que as percentagens são calculadas sobre 20 participantes e não sobre 21.

²²⁶ À data da realização do estudo.

respostas²²⁷ como se pode ver na figura 69. Seguiu-se o índice de assuntos da Memórias de África com quinze (15) respostas, o que situa a Memórias de África como aquela cujos SOC têm maior impacto no utilizador. Acresce referir que a Memórias de África foi o único repositório que não obteve respostas na opção “Nenhum”, o que parece indicar que este repositório tem cumprido o seu papel em termos de ajuda à pesquisa por assunto, ainda que possa fazer melhorias a nível da variedade de SOC disponíveis²²⁸. Em terceiro lugar ficou o RepositóriUM com treze (13) respostas a identificarem o SOC índice de assuntos. Os repositórios digitais que obtiveram maior número de respostas “Nenhum” foram a SinBAD com doze respostas (12), o Repositório Temático da Universidade do Porto com onze (11) respostas e o Repositório Aberto da Universidade do Porto com dez (10) respostas, o que parece indicar fraca visibilidade dos SOC destes repositórios²²⁹. Por último, na categoria “Nenhum”, o RepositóriUM obteve sete (7) respostas.

Questão 4 - “Qual o seu grau de satisfação com a utilização deste recurso de informação?”

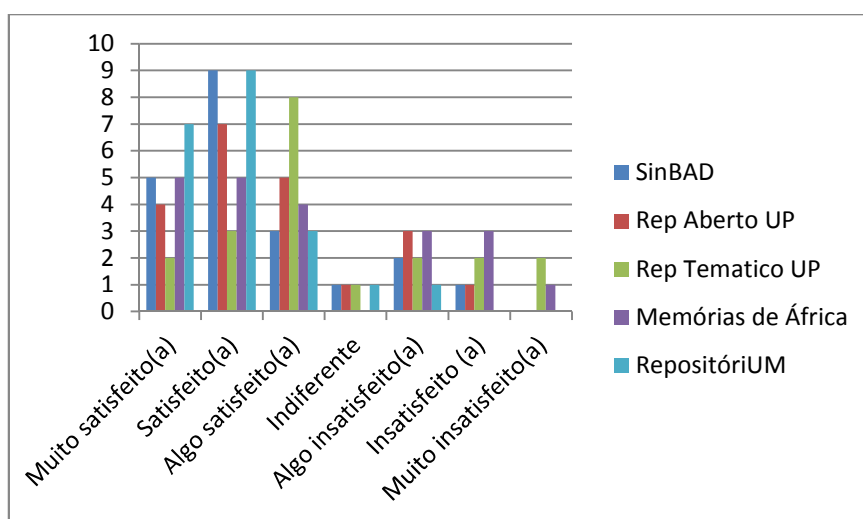


Figura 70 - Gráfico comparativo entre os diferentes repositórios digitais quanto ao grau de satisfação do utilizador

²²⁷ Dado que esta é uma questão de resposta múltipla, não se apresentam percentagens.

²²⁸ O facto de ter obtido classificação negativa na facilidade de acesso à informação merece discussão mais adiante, embora os motivos possam estar nalguns pormenores do *design* da IU por serem demasiado atrativos e distrativos para o utilizador e pelo posicionamento de um dos SOC.

²²⁹ Estes números representam mais de ou quase cinquenta por cento dos participantes, dado que quem respondeu “Nenhum” não assinalou mais nenhuma opção nesta questão.

Quanto ao grau de satisfação dos participantes no estudo para com os repositórios digitais avaliados (ver figura 70), o repositório digital que obteve maior número de respostas positivas nas respetivas três categorias - “muito satisfeito(a)”, “satisfeito(a)” e “algo satisfeito(a)” - foi o RepositóriUM com dezanove (19) respostas positivas (91%). Foi o mais votado com sete (7) respostas (33%) na categoria “Muito satisfeito(a)” e o mais votado na categoria “Satisfeito(a)”, *exaequo* com a SinBAD, com nove (9) respostas (43%). Em segundo lugar, ficou a SinBAD com dezassete (17) respostas positivas (81%), tendo sido o repositório digital mais votado, *exaequo* com o RepositóriUM, na categoria “Satisfeito(a)” com nove (9) respostas (43%), e o segundo mais votado (*exaequo* com a Memórias de África) na categoria “Muito satisfeito(a)” com cinco (5) respostas (24%). Foi, ainda, um dos três repositórios que não foi assinalado na resposta mais negativa – a categoria “muito insatisfeito(a)”.

A Memórias de África é o único repositório que não colhe indiferença dos participantes, mas, também, é aquele que obteve maior número de respostas negativas – categorias “algo insatisfeito(a)”, “insatisfeito(a)” e “muito insatisfeito(a)” – com sete (7) respostas. Na categoria “muito insatisfeito(a)” só dois (2) repositórios obtiveram respostas: o Repositório Temático da Universidade do Porto com duas (2) respostas e a Memórias de África com uma (1) resposta. O Repositório Temático da Universidade do Porto é o repositório com maior número de respostas negativas a seguir à Memórias de África – seis (6) respostas.

Questão 5 - “Considera que a presença de índices por assunto, ontologias, tesauros, classificações, taxonomias, etc., facilita a realização da pesquisa por assunto?”

Quanto à perceção, pelos participantes, do papel dos SOC como instrumentos de apoio à pesquisa por assunto, as respostas foram, totalmente, unânimes no sentido afirmativo conforme se pode verificar na figura 71. Todos os participantes assinalaram a resposta “sim” a esta questão, nos cinco questionários posteriores (QP1 a QP5).

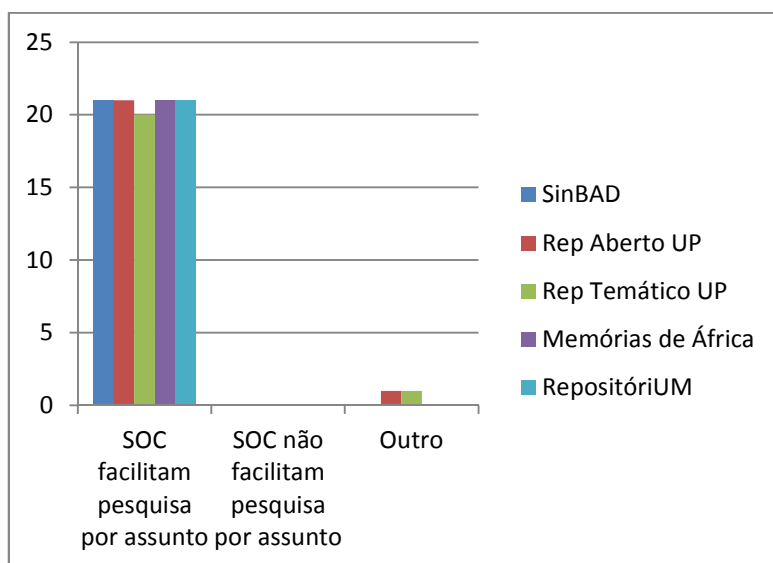


Figura 71 - Gráfico comparativo entre os diferentes repositórios digitais quanto ao papel facilitador da presença dos SOC na pesquisa por assunto

Houve duas respostas abertas nos questionários relativos aos repositórios da Universidade do Porto, submetidas pelo mesmo participante, nas quais se referia que “facilita se o índice estiver relacionado com os conteúdos” (Repositório Aberto da Universidade do Porto) e que “facilita se estiver relacionado com o documento, o que nem sempre aconteceu” (Repositório Temático da Universidade do Porto). Estas respostas podem ser interpretadas como uma crítica à análise e representação da informação contida naqueles repositórios digitais (e recuperada na pesquisa efetuada nos testes 2 e 3) e não, propriamente, aos SOC presentes nos mesmos.

4.1.2. Testes de usabilidade com *eyetracking*

Os exemplos abaixo apresentados são, na sua quase totalidade, extraídos do teste 1 (SinBAD) por ter sido aquele que reuniu dados mais completos e complexos, conseguindo ter uma página em que há fixações de todos os participantes no estudo, pelo que permite efetuar o maior número de relações entre variáveis.

Os mapas de fixações revelam padrões de fixação em determinados grupos, indicando um comportamento informacional comum aos participantes com determinadas características comuns.

Foi possível detetar um padrão de maior concentração das fixações do olhares dos participantes nas funcionalidades de recuperação da informação que permitem atingir o objetivo da pesquisa em dois grupos de participantes: os que pertencem à categoria “docente e investigador” e os que pertencem à categoria “profissional de informação”. Esse padrão é comum ao grupo de participantes que compõe o grupo etário “40 a 49 anos”, o que não surpreende – os participantes de 40 a 49 anos (ver figura 72) pertencem na sua maioria àqueles dois grupos de atividade na Universidade de Aveiro.

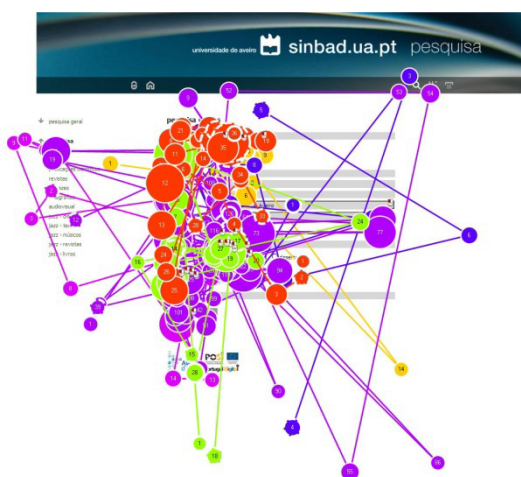


Figura 72 - Mapa de fixações dos participantes com 40-49 anos (Teste 1)

O grupo cujo padrão indica menor dispersão ou distração durante a pesquisa é o resultante da relação entre 3 categorias ou variáveis: docente ou investigador com doutoramento e com 40 a 49 anos (ver figura 73). O diâmetro das fixações e o tempo de fixação confirmam esta análise.



Figura 73 - Mapa de fixações dos participantes com doutoramento, 40-49 anos, e docentes ou investigadores (Teste 1)

Porém, um outro padrão, igualmente focado, mas, dividido pela pesquisa simples (preferencial no caso anterior), pelo campo de assunto, pelo tipo de documento e pelo texto livre, foi detetado na categoria “ensino secundário (12º ano)” (ver figura 74). Dado que essa categoria se aplica a um único participante que pertence à categoria “profissional de informação” e ao grupo etário 30-39 anos, parece tratar-se de um caso em que a atividade profissional²³⁰ tem influência no comportamento informacional.

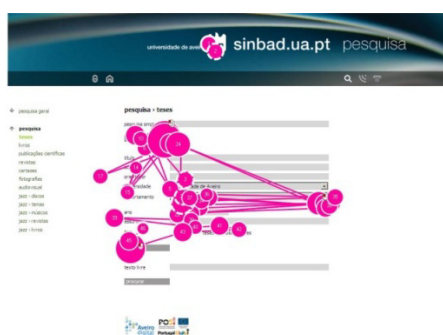


Figura 74 – Mapa de fixações dos participantes²³¹ com habilitações académicas ao nível do ensino secundário (Teste 1)

²³⁰ E, possivelmente, a formação profissional, dado que ao nível do ensino secundário há cursos da área da informação, e dado que o participante pode ter frequência de cursos de formação profissional no âmbito da sua atividade profissional, promovidos ou financiados pela própria organização em que se insere.

²³¹ Pertencente ao grupo etário 30-39 anos e profissional de informação – um participante.

Os participantes da categoria “mestrado” (ver figura 75) relativa a habilitações académicas, apresentaram um padrão comum de distração com pormenores de layout do *design* da IU, assim como os do grupo etário “30-39 anos”.

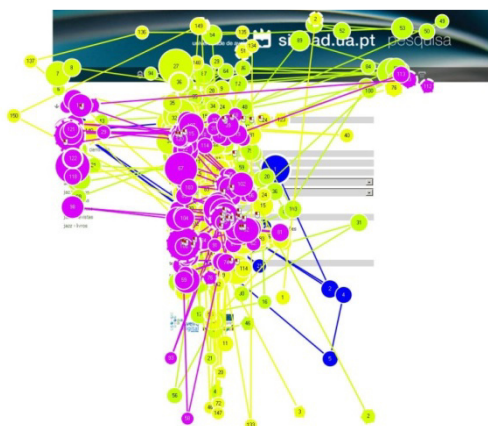


Figura 75 - Mapa de fixações dos participantes com o grau de mestre (Teste 1)

No grupo de participantes da categoria “licenciatura” (ver figura 77) há um padrão comum que revela alguma distração com pormenores de *layout* do *design* da IU, mas, em menor grau do que o da categoria “mestre”. Excetua-se o caso de um participante, cujo mapa de fixações parece indiciar um caso de *user lostness*, em consonância com os dados recolhidos durante o *think-aloud* (ver figura 76).

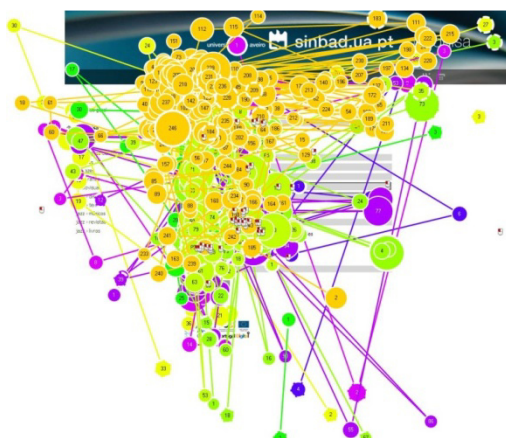


Figura 76 - Mapa de fixações dos participantes com o grau de licenciado (Teste 1)

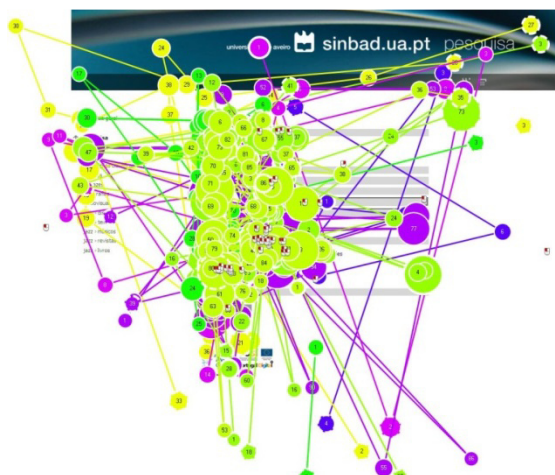


Figura 77 - Mapa de fixações dos participantes com o grau de licenciado, excluindo o ID 19²³² (Teste 1)

O mesmo se verifica em todos os grupos a que o mesmo participante pertence como o grupo etário “18-29 anos” (ver figura 79) e a categoria de atividade “aluno de licenciatura/mestrado integrado”. O padrão diverge do dos restantes membros do grupo (ver figuras 76 e 78).

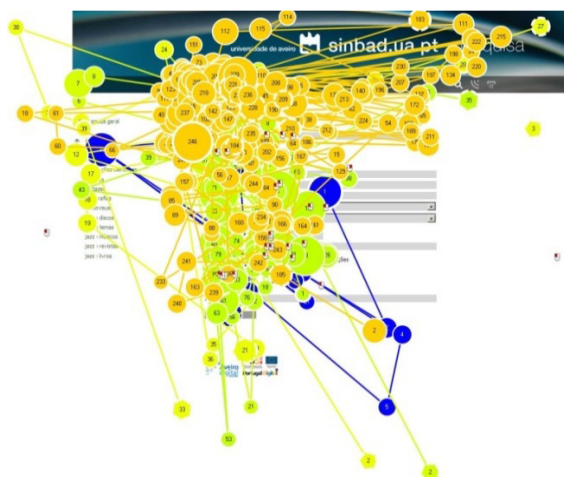


Figura 78 - Mapa de fixações dos participantes com 18-29 anos (Teste 1)

²³² Utilizador com comportamento típico de “user lostness”.

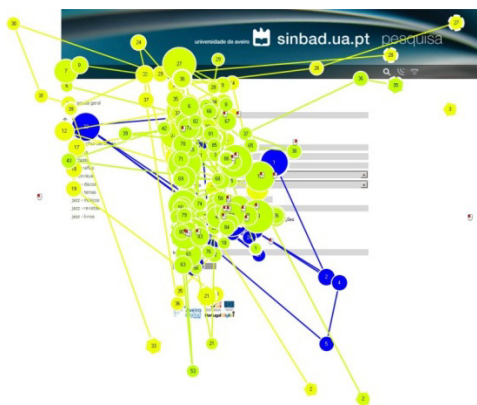


Figura 79 - Mapa de fixações dos participantes com 18-29 anos, excluindo o participante com o ID 19 (Teste 1)

O mapa de fixações do conjunto de participantes que não conseguiram identificar nenhum SOC (ver figura 80) é, igualmente, pouco focado, possivelmente, devido ao repetido rastreo visual da IU com o olhar, o que é um movimento natural quando não é repetido insistentemente. Neste último caso, poderá denotar alguma dificuldade na localização de elementos do SRI na IU.



Figura 80 - Mapa de fixações dos participantes que não identificaram nenhum SOC (Teste 1)

A seguir apresentam-se exemplos de mapas de fixações dos outros quatro repositórios digitais avaliados, por ordem de realização dos testes de usabilidade com *eyetracking* no sistema QUETTA. Abaixo, na figura 81, segue um exemplo de visualização da IU do Repositório Aberto da Universidade do Porto.

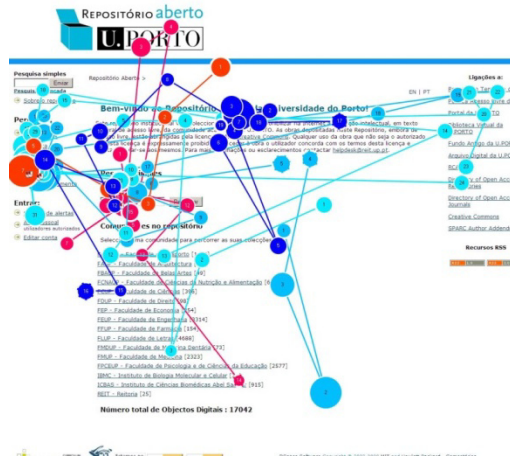


Figura 81 - Repositório Aberto da Universidade do Porto (Teste 2)

Depois, na figura 82, um exemplo de visualização da IU do Repositório Temático da Universidade do Porto que denota alguma repetição de movimentos oculares.

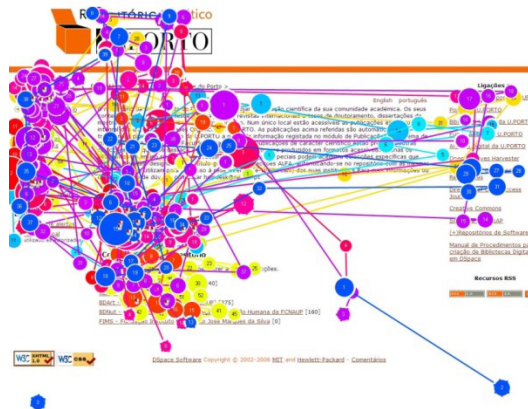


Figura 82 - Repositório Temático da Universidade do Porto (Teste 3)

No caso da Memórias de África mantida pela Universidade de Aveiro inserem-se abaixo três exemplos de visualização da IU, nas figuras 83, 84 e 85, dado que a variedade de documentos digitais, inclusive no cabeçalho²³³ da *webpage*, atrai o olhar do participante, bem como o índice de assuntos.

²³³ Sempre em rotação.

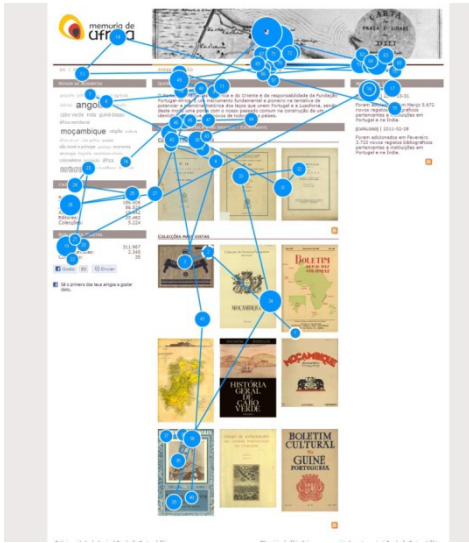


Figura 83 - Memórias de África mantida pela Universidade de Aveiro (Teste 4) – exemplo 1

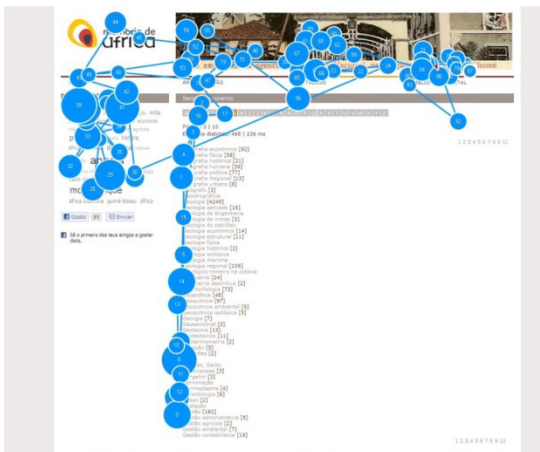


Figura 84 - Memórias de África mantida pela Universidade de Aveiro (Teste 4) – exemplo 2

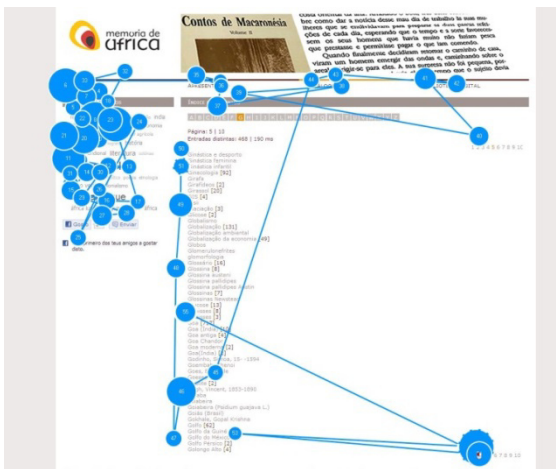


Figura 85 - Memórias de África mantida pela Universidade de Aveiro (Teste 4) – exemplo 3

Por último, na figura 86, fica um exemplo de visualização da IU do RepositóriUM da Universidade do Minho.

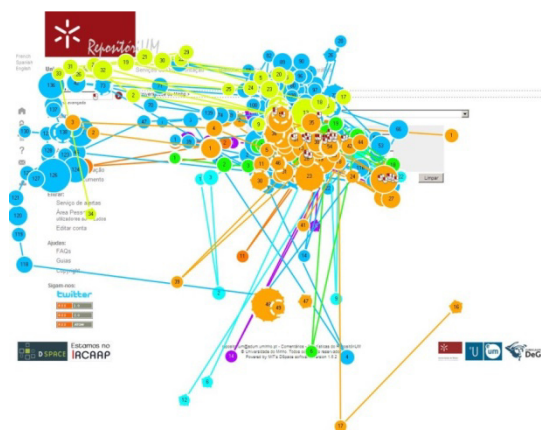


Figura 86 - RepositóriUM da Universidade do Minho (Teste 5)

4.1.3. *Think-aloud*

No *think-aloud* foram realizadas anotações dos tempos de execução das tarefas e das observações verbais e não-verbais voluntárias dos participantes.

O gráfico da figura 87 elucida acerca da diferença entre a duração mínima e a duração máxima de cada sessão do sistema QUETTA que oscilou entre os treze (13) minutos e os sessenta e três (63) minutos. A maioria (86%) das sessões situou-se entre os 18 e os 34 minutos de duração do estudo com o sistema QUETTA, desde o início do questionário prévio até ao fim do questionário posterior 5.

No gráfico da figura 88, são apresentados os tempos parciais de cada teste por participante com maior pormenor, permitindo perceber que houve dois casos de tempos extremos nos dois últimos utilizadores e que não se deveram a problemas do equipamento ou a perturbação da sala com ruído. Um dos participantes, o que demorou 63 minutos é o caso de *user lostness*, e o outro, o que demorou 45 minutos, esteve demoradamente a analisar a informação recuperada numa pesquisa.

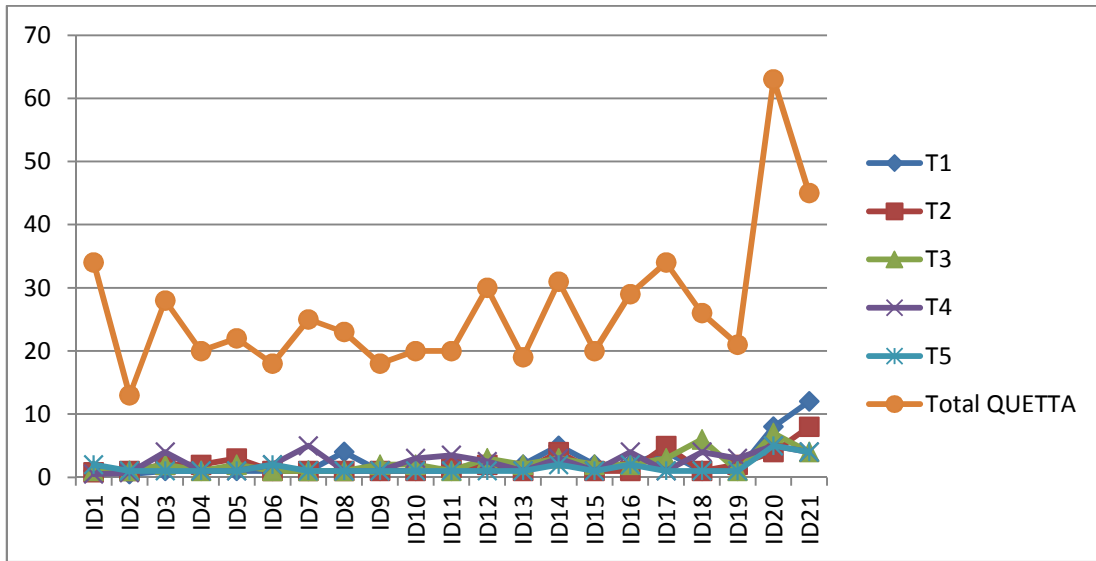


Figura 87 – Gráfico dos tempos, em minutos, de realização dos testes de usabilidade com *eyetracking* 1 a 5, excluindo os tempos de calibragem 1 a 5 e de instruções para o teste 1 a 5, mais o tempo total de realização do estudo com o sistema QUETTA por participante.

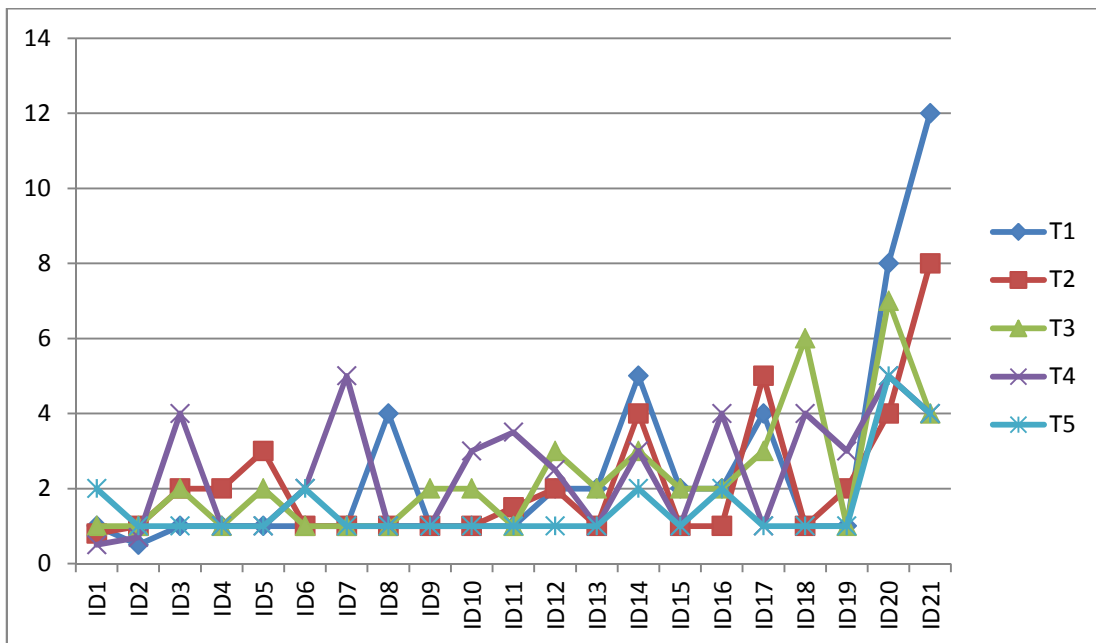


Figura 88 - - Gráfico dos tempos, em minutos, de realização dos testes de usabilidade com *eyetracking* 1 a 5, excluindo os tempos de calibragem 1 a 5 e de instruções para o teste 1 a 5, por participante.

Na generalidade, não houve muita expressão verbal e não-verbal ao longo das várias sessões, tendo havido algumas pesquisas sem sucesso e/ou abortadas.

4.2. Discussão dos resultados

Neste ponto, proceder-se-á a uma discussão mais alargada dos resultados já descritos no ponto 4.1., pelo que será inevitável repescar ideias já enunciadas para que seja compreensível a sequência das ilações retiradas. Os dados analisados no início deste capítulo, permitem várias ilações e conclusões.

No capítulo 1, nas páginas 69 e 70, foi esclarecido que este projeto de investigação, ao centrar-se na perspetiva do utilizador, utiliza os seguintes critérios de avaliação da usabilidade:

- ✚ Ajuda e recuperação de erros
- ✚ Atingir objetivos
- ✚ Eficiência
- ✚ Satisfação
- ✚ *User lostness*

A **ajuda e recuperação de erros** é um critério definido por Blanford e Buchanan em 2003. De entre os dados recolhidos e tratados, é possível identificar os resultados que permitem a utilização deste critério na avaliação da usabilidade dos repositórios digitais abrangidos por este estudo. As respostas à questão 2 dos questionários posteriores 1, 2, 3, 4 e 5 fornecem os dados necessários para o efeito. Segue-se a tabela com os dados utilizados para criar o gráfico da figura 68 que consta na página 180 desta tese.

Hipóteses	SinBAD	Rep Aberto UP	Rep Tematico UP	Memórias de África	RepositóriUM
Encontrou auxílio na pesquisa por assunto	11	13	6	11	11

Não encontrou auxílio na pesquisa por assunto	10	8	14	10	10
--	----	---	----	----	----

No que diz respeito à ajuda, o repositório melhor classificado foi o Repositório Aberto da Universidade do Porto (UP), com 62% de respostas positivas (13 respostas). O pior classificado foi o Repositório Temático da UP com, apenas, 30% de respostas positivas (6), o que constitui um desempenho negativo. Os restantes repositórios digitais obtiveram uma classificação positiva, com 52% de respostas positivas (11). No que diz respeito à recuperação de erros, na generalidade dos participantes não se observaram erros. Todavia, foi perceptível que o Repositório Temático da UP causou maior número de hesitações.

Atingir objetivos é outro critério definido por Blanford e Buchanan em 2003. As respostas à questão 5 dos questionários posteriores 1, 2, 3, 4 e 5 fornecem os dados que permitem a utilização deste critério. Segue-se a tabela com os dados utilizados para criar o gráfico da figura 71 que consta na página 184 desta tese.

Hipóteses	SinBAD	Rep Aberto UP	Rep Temático UP	Memórias de África	RepositóriUM
SOC facilitam pesquisa por assunto	21	21	20	21	21
SOC não facilitam pesquisa por assunto	0	0	0	0	0
Outro	0	1	1	0	0

Na sua essência, os SOC têm por objetivo serem facilitadores da pesquisa por assunto, pelo que a formulação da questão aponta para o critério de avaliação da usabilidade “Atingir objetivos”. O desempenho dos cinco repositórios digitais neste critério é francamente positivo, tendo sido de 100% em todos os casos. A linha “Outro”, na tabela, refere-se a respostas abertas que, nos casos em concreto, continham observações relacionadas com a indexação dos documentos contidos nos repositórios da Universidade do Porto.

A **eficiência** é um critério definido por Nielsen em 1993, pela ISO em 1998 e por Oulanov e Pajarillo em 2002. Utiliza-se o critério na definição dada por Nielsen: a rapidez com que o utilizador leva a cabo as suas tarefas.

A figura 88, constitui a melhor ilustração deste critério. O RepositoriUM destaca-se no gráfico como sendo o repositório digital em que o tempo de execução da pesquisa é mais reduzido na generalidade dos participantes – é, seguramente, o mais eficiente dos cinco repositórios avaliados. Segue-se a tabela com os dados utilizados para criar o gráfico da figura 88 que consta na página 193 desta tese.

Participante	T1	T2	T3	T4	T5
ID1	1	0,8	1	0,5	2
ID2	0,5	1	1	0,7	1
ID3	1	2	2	4	1
ID4	1	2	1	1	1
ID5	1	3	2	1	1
ID6	1	1	1	2	2
ID7	1	1	1	5	1
ID8	4	1	1	1	1
ID9	1	1	2	1	1
ID10	1	1	2	3	1
ID11	1	1,5	1	3,5	1
ID12	2	2	3	2,5	1
ID13	2	1	2	1	1
ID14	5	4	3	3	2
ID15	2	1	2	1	1
ID16	2	1	2	4	2
ID17	4	5	3	1	1
ID18	1	1	6	4	1
ID19	1	2	1	3	1
ID20	8	4	7	5	5
ID21	12	8	4	4	4

A **satisfação** é outro critério definido por Nielsen em 1993 e pela ISO em 1998. As respostas à questão 4 dos questionários posteriores 1, 2, 3, 4 e 5 fornecem os dados a serem utilizados para o efeito. Segue-se a tabela com os dados utilizados para criar o gráfico da figura 70 que consta na página 182 desta tese.

Grau de satisfação / insatisfação	SinBAD	Rep Aberto UP	Rep Temático UP	Memórias de África	RepositóriUM
Muito satisfeito(a)	5	4	2	5	7
Satisfeito(a)	9	7	3	5	9
Algo satisfeito(a)	3	5	8	4	3
Indiferente	1	1	1	0	1
Algo insatisfeito(a)	2	3	2	3	1
Insatisfeito (a)	1	1	2	3	0
Muito insatisfeito(a)	0	0	2	1	0

O repositório digital que deixa o utilizador mais satisfeito é o RepositóriUM, dado que obteve 91% de respostas positivas nas três categorias relativas a grau de satisfação (“muito satisfeito(a)”, “satisfeito(a)” e “algo satisfeito(a)”), ou seja, a percentagem mais elevada. O RepositóriUM foi o mais votado com 33% na categoria “Muito satisfeito(a)”, e o mais votado na categoria “Satisfeito(a)”, *exaequo* com a SinBAD, com 43%. Em segundo lugar, ficou a SinBAD com 81%, tendo sido o repositório digital mais votado, *exaequo* com o RepositóriUM, na categoria “Satisfeito(a)” com 43%, e o segundo mais votado (*exaequo* com a Memórias de África) na categoria “Muito satisfeito(a)” com 24%. Foi, ainda, um dos três repositórios que não foi assinalado na resposta mais negativa – a categoria “muito insatisfeito(a)”.

A Memórias de África é o repositório digital que obteve maior número de respostas negativas (categorias “algo insatisfeito(a)”, “insatisfeito(a)” e “muito insatisfeito(a)”) com sete (7) respostas. Na categoria “muito insatisfeito(a)” só dois (2) repositórios obtiveram respostas: o Repositório Temático da Universidade do Porto com duas (2) respostas e a Memórias de África com uma (1) resposta. O Repositório Temático da Universidade do Porto é o repositório com maior número de respostas negativas a seguir à Memórias de África – seis (6) respostas.

Em termos absolutos, o RepositóriUM, a SinBAD e o Repositório Aberto foram considerados os repositórios mais satisfatórios, mas nenhum dos cinco foi considerado maioritariamente insatisfatório, pelo que os resultados são globalmente positivos.

User lostness é um critério, unicamente, definido por Nielsen em 1993: o utilizador não consegue identificar onde se encontra dentro da IU, é incapaz de regressar a informação anteriormente visitada, não consegue direcionar-se para informação que se acredita existir ou não consegue recordar-se dos pontos-chave. O mapa de fixações constante da figura 78 identifica um caso provável de um utilizador que se sentiu perdido na IU da SinBAD. Se se confrontarem estes dados com os da figura 88, depreende-se que o participante em causa demorou mais tempo a concretizar as pesquisas do que os outros participantes no estudo em todas as IU. Um dos motivos para a demora poderá ser o número de fixações repetidas. Fica a interrogação se essas fixações resultam de o participante ficar perdido nas IU e deambular pelas mesmas sem conseguir concretizar com eficiência a pesquisa, ou se resultam de movimentos oculares involuntários de rastreio visual. E se a maior demora na concretização das pesquisas evidencia um menor conhecimento de repositórios digitais e perda de tempo a explorar essas plataformas. Cruzando os dados com as respostas aos questionários, conclui-se que a demora poderá dever-se a menor experiência e conhecimento destas plataformas digitais.

As respostas às quatro questões que desencadearam este projeto de investigação são agora passíveis de serem respondidas, pelo que são aqui apresentadas, bem como a confirmação ou a negação²³⁴ das hipóteses formuladas no início da investigação²³⁵.

4.2.1. Acessibilidade dos SOC nos SI digitais do tipo arquivo digital e do tipo biblioteca digital

Como atrás ficou dito, no ponto anterior deste capítulo, os participantes no estudo com o sistema QUETTA identificaram vários SOC nos repositórios digitais em avaliação. Porém, nem em todos os casos os SOC foram bem identificados, havendo, mesmo, SOC inexistentes nos repositórios digitais que foram identificados por

²³⁴ Todas as hipóteses foram confirmadas, pelo que não houve negação.

²³⁵ Cf. p. 26-27.

alguns participantes²³⁶. Pela observação feita, em simultâneo, com o *think-aloud*, tornou-se evidente o desconhecimento de alguns dos conceitos associados à terminologia apresentada no questionário por parte de alguns dos participantes. Os problemas com terminologias específicas de grupos profissionais são recorrentes, mas esse facto pode, igualmente, denotar que os SOC não têm tido divulgação suficiente no âmbito dos repositórios digitais. Haverá quem argumente que não são necessários, mas este estudo demonstra uma perceção diversa do utilizador, mesmo quando nem sempre os sabe identificar bem – a importância e a necessidade dos SOC como instrumentos de apoio à pesquisa por assunto é reconhecida. O protocolo *think-aloud* foi muito importante, ainda que a expressividade dos participantes tenha sido discreta, para perceber as opiniões dos participantes. Embora tenham sido disponibilizadas questões de resposta aberta nos questionários, alguns participantes revelaram perceções importantes no final da sessão, ou durante a mesma, em voz audível. A pouca diversidade de SOC nos repositórios digitais foi notada, bem como a pouca presença ou ausência de SOC de contextualização semântica como os tesouros.

De entre as questões que deram origem a este projeto de investigação, a questão 1 indagava acerca da acessibilidade dos SOC nos SI digitais do tipo arquivo digital e biblioteca digital, da seguinte forma: “Os SOC quando presentes nos SI digitais do tipo arquivo ou biblioteca estão acessíveis?” Na sequência dessa questão foi formulado uma hipótese – a hipótese 1 – que carece de confirmação e que foi plasmada na seguinte frase: “Os SOC encontram-se acessíveis na página inicial do SI digital”. Após a análise dos dados recolhidos com o sistema QUETTA, **é possível confirmar esta hipótese**. Conclui-se, portanto, que os sistemas de organização do conhecimento se encontram acessíveis na página inicial dos sistemas de informação digitais avaliados.

²³⁶ À data da realização do estudo.

4.2.2. Os SOC como facilitadores da recuperação da informação por assunto

Para quem é da área profissional ou académica da informação, constitui uma evidência que os SOC sejam facilitadores da recuperação da informação por assunto, dado que constituem instrumentos de apoio à pesquisa. Todavia, quem contacta com o meio académico de outras áreas científicas nem sempre percebe que essa constatação seja um dado adquirido. Por isso, o resultado deste estudo foi algo surpreendente, até porque incluiu participantes de outras áreas científicas como a engenharia informática e a engenharia eletrónica e de telecomunicações.

A resposta dos participantes foi unânime – todos consideram que os SOC facilitam a recuperação da informação por assunto. Portanto, a resposta à questão 2 “A utilização de SOC nos SI digitais facilita a recuperação da informação por assunto?” é afirmativa, em absoluto, pelo que **a hipótese 2 formulada** na frase “A utilização de SOC facilita a recuperação da informação por assunto no SI digital” **se confirma**.

4.2.3. A influência positiva do conhecimento prévio dos SOC pelo utilizador como fator de sucesso e de eficiência na pesquisa por assunto

Mediante o cruzamento de dados, nomeadamente, dos participantes que costumam pesquisar mais por autor e título, e, sobretudo, dos que não identificaram nenhum SOC nos repositórios digitais em avaliação, e até dos que identificaram SOC que não existiam no repositório digital em causa, foi possível concluir que o conhecimento prévio dos SOC influencia positivamente a pesquisa por assunto. As pesquisas sem sucesso e os tempos de execução de duração mais longa coincidem com aqueles perfis.

A resposta à questão 3 – “O conhecimento prévio dos SOC por parte do utilizador influencia positivamente o sucesso na pesquisa por assunto, diminuindo o tempo de

obtenção de resultados satisfatórios?” é, também, afirmativa. Assim, **a hipótese 3** – “O conhecimento prévio dos SOC por parte do utilizador influencia positivamente o sucesso na pesquisa por assunto, diminuindo o tempo de obtenção de resultados satisfatórios” – **fica confirmada** pelos resultados deste projeto de investigação.

4.2.4. A influência da presença de SOC no SI digital no grau de satisfação do utilizador para com o mesmo

Os dados recolhidos e analisados permitem responder afirmativamente à última questão que impulsionou este projeto de investigação, ainda que se possa ver noutros fatores a causa do grau de satisfação do utilizador. O grau de satisfação dos participantes foi, predominantemente, positivo. Há algumas respostas nos questionários que poderiam parecer incompreensíveis sem a combinação daqueles com as outras duas técnicas de investigação do sistema QUETTA, e sem os resultados destas, a resposta a esta questão poderia ser menos segura. Na realidade, foi o cruzamento de dados recolhidos nos questionários com os mapas de fixações, os vídeos do *eyetracker* e as anotações do *think-aloud* que permitiu esclarecer dúvidas quanto à classificação negativa atribuída por 33% dos participantes no estudo ao repositório digital Memórias de África, dado que entre os repositórios avaliados é um dos que mais investe nos SOC e aquele que obteve maior número de identificações unívocas com a nuvem de assuntos e o índice de assuntos. O que se detetou no *think-aloud* e na análise das fixações, é que o posicionamento do *link* para aceder ao índice não era o mais perceptível²³⁷ e que o *layout* do *design* da IU por ser atrativo constitui um fator de distração do utilizador. O cabeçalho dinâmico com imagens de documentos digitais em rotação contínua dispersa o olhar do utilizador.

O facto de que o grau de satisfação com os repositórios digitais avaliados oscilou entre os 67% e os 95%, e o cruzamento desses dados dos resultados da questão 4 dos questionários posteriores 1 a 5 com os dos resultados das questões 3 e 5 dos mesmos

²³⁷ Esse problema foi, mais tarde, analisado pela equipa que gere o repositório digital, tendo a mesma efetuado alterações na organização dos conteúdos na IU.

questionários, permite concluir que a resposta à questão 4 – “A presença dos SOC nos SI digitais influencia o grau de satisfação do utilizador com o SI digital?” é positiva. **A hipótese 4** formulada na frase “A presença dos SOC influencia positivamente o grau de satisfação do utilizador com o SI digital” **foi confirmada pelos resultados deste projeto de investigação.**

Os resultados desta investigação realçam, pois, o papel dos sistemas de organização do conhecimento na recuperação da informação e na vida quotidiana de milhões de pessoas que acedem a plataformas digitais da mais variada tipologia com o objetivo de satisfazerem necessidades de informação, mais ou menos prementes, cuja satisfação pode ter um considerável impacto na vida académica, profissional e particular de cada cidadão.

Conclusão

1. Considerações gerais

A primeira conclusão a retirar dos resultados deste projeto de investigação é que os sistemas de organização do conhecimento são reconhecidos como ferramentas importantes na recuperação da informação quer pelos profissionais de informação quer por utilizadores dos mais variados contextos demográficos, socioprofissionais, académicos e informacionais²³⁸. A relevância da pesquisa e recuperação da informação no dia-a-dia de cidadãos de todo o mundo e a omnipresença das redes de comunicação de dados e das plataformas digitais na sua vida, em função de uma sociedade da informação e do conhecimento, tem vindo a alargar a penetração do mundo digital em todas as formas de organização social mediante o comércio eletrónico, o governo eletrónico, as redes sociais e as redes de conhecimento. O cidadão comum acede, diariamente, a plataformas digitais com os mais variados objetivos e, entre eles, a necessidade de pesquisar e recuperar informação. Se os sistemas de organização do conhecimento são reconhecidos como sendo um importante apoio neste processo, será lógico investir mais no seu desenvolvimento e aperfeiçoamento e na diversificação da sua presença nas plataformas digitais. A falta de SOC de contextualização semântica e a pouca diversidade de SOC nos

²³⁸ O contexto informacional dos participantes no estudo não era discrepante, porém, também, não era uniforme.

repositórios digitais avaliados pelo estudo realizado com o sistema QUETTA neste projeto de investigação, reforça a necessidade de um maior investimento neste tipo de ferramentas.

Conclui-se, igualmente, que é necessário colocar SOC nas plataformas digitais, mas com menus de ajuda e com apoio ao utilizador. A formação para o utilizador ministrada, e bem, em muitos serviços de informação, não tem penetração no mundo digital a menos que comece a ser feita *online* em *webinars*, por exemplo. Mas, isso não será suficiente, dado que a Internet está disponível ao longo de 365 dias no ano, durante 24 horas por dia, em qualquer dispositivo e, cada vez mais, nos dispositivos móveis. Quem pesquisa, quer tirar as dúvidas que tiver naquele momento, se as tiver, ou melhor, se perceber que precisa de auxílio. É evidente que o ideal será ter SOC intuitivos que dispensem muita formação e muitos esclarecimentos. Para lá chegarmos, serão necessárias equipas multidisciplinares de cientistas da informação, de linguistas, de engenheiros informáticos, de cientistas da computação, de neurolinguistas, e outros, eventualmente.

2. Limitações

O projeto de investigação atrás descrito teve algumas limitações no seu decurso que alteraram a sua calendarização e adiaram a sua conclusão, sobretudo, as relacionadas com questões de financiamento.

A indisponibilidade de um laboratório de usabilidade equipado com um *eyetracker* em permanência foi a primeira limitação sentida, mas, acabou por ser ultrapassada da forma que já foi referida ao longo desta tese.

O limitado intervalo de tempo em que houve disponibilidade do *eyetracker*, cedido pela Portugal Telecom ao CETAC que, por sua vez, o cedeu a este projeto de investigação, não permitiu o reagendamento das sessões de dois voluntários que faltaram e que se disponibilizaram para realizarem as sessões nas semanas seguintes, nem permitiria, se tivesse sido necessário, a repetição dos testes de usabilidade com *eyetracking* em tempo útil.

Por último, a falta de tempo disponível para analisar todas as potenciais relações a extrair da imensa quantidade de dados recolhida com o sistema QUETTA no intervalo de tempo do doutoramento.

3. Investigação futura

Este projeto de investigação abriu perspectivas de investigação e desenvolvimento para os próximos anos.

A primeira investigação a desenvolver no futuro será a aplicação do sistema QUETTA a outros repositórios digitais e em condições laboratoriais que permitam um trabalho de investigação continuado. A avaliação de outros sistemas de informação com o modelo integrado de estudo do comportamento do utilizador concebido neste projeto de investigação, permitirá uma análise comparativa mais alargada de vários SI do tipo arquivo digital e biblioteca digital.

Outra investigação a desenvolver nos próximos anos poderá ser o alargamento do sistema QUETTA com a inserção de outra(s) vertente(s) de análise. Haverá que ponderar acerca da fiabilidade de uma ferramenta de leitura das expressões faciais e da possibilidade da sua integração no QUETTA, por exemplo, bem como da sua complementaridade.

4. Contributos

O principal contributo deste projeto de investigação foi a conceção de um modelo integrado de estudo do comportamento do utilizador – o sistema QUETTA – que poderá ser aplicado a outros sistemas de informação digitais.

Outro contributo deste projeto foi a avaliação da presença dos sistemas de organização do conhecimento nos repositórios digitais de três instituições de ensino superior em Portugal: a Universidade do Porto, a Universidade de Aveiro e a Universidade do Minho.

O último contributo decorre dos resultados da investigação obtidos - o reconhecimento do importante papel dos sistemas de organização do conhecimento nos repositórios digitais e o seu impacto na sociedade da informação e do conhecimento, bem como a necessidade de os diversificar e de os melhorar aumentando a sua presença naqueles sistemas de informação digitais.

Referências bibliográficas

ADAMS, A. and BLANDFORD, A., 2002. Acceptability of medical digital libraries. In: *Health Informatics Journal* [online]. 1 June 2002. Vol. 8, no. 2, pp. 58–66. [Accessed 2 December 2013]. DOI 10.1177/146045820200800202. Available from: <http://jhi.sagepub.com/content/8/2/58.refs>.

ALBERT, Bill, TULLIS, Tom and TEDESCO, Donna, 2010. *Beyond the usability lab: conducting large-scale online user experience studies*. [1st]. Burlington: Morgan Kaufmann Publishers. ISBN 978-0-12-374892-8.

ARMS, William Y., HILLMANN, Diane, LAGOZE, Carl, KRAFFT, Dean, MARISA, Richard, SAYLOR, John, TERRIZZI, Carol and VAN DE SOMPEL, Herbert, 2002. A Spectrum of Interoperability: The Site for Science Prototype for the NSDL. In: *D-Lib Magazine* [online]. 2002. Vol. 8, no. 1. [Accessed 6 December 2013]. DOI 10.1045/january2002-arms. Available from: <http://www.dlib.org/dlib/january02/arms/01arms.html>.

BAEZA-YATES, RICARDO RIBEIRO-NETO, Berthier, 1999. *Modern information retrieval*. New York: ACM Press. ISBN ISBN 0-201-39829-X.

BELL, W. J., 1991. *Searching behavior: the behavioral ecology of finding resources*. New York: Chapman & Hall.

BERTOT, John Carlo [et al.], 2006. Functionality, usability, and accessibility: iterative user-centered evaluation strategies for digital libraries. In: *Performance Measurement and Metrics*. 2006. Vol. 7, no. 1, pp. 17–28.

BEYNON-DAVIES, Paul, 2013. *Business Information Systems* [online]. 2nd ed. S.l.: Palgrave Macmillan. ISBN 9781137265807. Available from: <http://www.palgrave.com/products/title.aspx?pid=569375>.

BEYNON-DAVIES, P., 2002. *Information Systems: An introduction to Informatics in Organisations* 1st ed., Palgrave Macmillan. ISBN 0333963903.

BEYNON-DAVIES, P., 2009. *Business Information Systems* 1st ed., Palgrave Macmillan. ISBN 9780230203686.

BISHOP, Ann Peterson, 1998. Logins and Bailouts: Measuring Access, Use, and Success in Digital Libraries. In: [online]. 30 December 1998. [Accessed 30 November 2013]. Available from: <http://quod.lib.umich.edu/jjep/3336451.0004.207?rgn=main;view=fulltext>.

BISHOP, Ann Peterson, NEUMANN, Laura J., STAR, Susan Leigh, MERKEL, Cecelia, IGNACIO, Emily and SANDUSKY, Robert J., 2000. Digital libraries: Situating use in changing information infrastructure. In: *Journal of the American Society for Information Science* [online]. 2000. Vol. 51, no. 4, pp. 394–413. [Accessed 10 November 2013]. DOI 10.1002/(SICI)1097-4571(2000)51:4<394::AID-ASI8>3.0.CO;2-Q. Available from: [http://doi.wiley.com/10.1002/\(SICI\)1097-4571\(2000\)51:4<394::AID-ASI8>3.0.CO;2-Q](http://doi.wiley.com/10.1002/(SICI)1097-4571(2000)51:4<394::AID-ASI8>3.0.CO;2-Q).

BLANDFORD, A., KEITH, S., CONNELL, I. and EDWARDS, H., 2004. Analytical usability evaluation for digital libraries: a case study. In: *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*. S.l.: s.n. 2004. pp. 27–36.

BLANFORD, Ann and BUCHANAN, George, 2003. Usability of digital libraries: a source of creative tensions with technical developments. In: *IEEE-CS Technical Committee on Digital Libraries' on-line newsletter* [online]. 2003. Available from: <http://www.ieee-tcdl.org/Bulletin/v1n1/blandford/blandford.html>.

BOLLEN, J. and LUCE, R., 2002. Evaluation of Digital Library Impact and User Communities by Analysis of Usage Patterns. In: *D-Lib Magazine* [online]. 2002. Vol. 8, no. 6. Available from: <http://www.dlib.org/dlib/june02/bollen/06bollen.html>.

BOLLEN, J., VEMULAPALLI, S. AND XU, W., 2002. Digital Library Evaluation by Analysis of User Retrieval Patterns. In M. Agosti & C. Thanos, eds. *6th European Conference on Digital Libraries*. Springer-Verlag, pp. 161–171.

BORGES, Maria Manuel, 2007. *A Esfera: Comunicação Académica e Novos Media* [online]. 7 May 2007. S.l.: s.n. [Accessed 20 November 2013]. Available from: <https://estudogeral.sib.uc.pt/jspui/handle/10316/8557>.

Tese de doutoramento em Letras, área de Ciências Documentais (Tecnologia da Informação) apresentada à Fac. de Letras de Coimbra

BORGMAN, Christine L., 1999. What are digital libraries? Competing visions. In: *Information Processing & Management*. 1999. Vol. 35, no. 3, pp. 227–243.

BRINCK, Tom, GERGLE, Darren and WOOD, Scott D., 2001. *Usability for the Web: Designing Web Sites that Work (Interactive Technologies)* [online]. S.l.: Morgan Kaufmann. [Accessed 12 November 2013]. ISBN 1558606580. Available from: <http://www.amazon.com/Usability-Web-Designing-Interactive-Technologies/dp/1558606580>.

BROPHY, Peter and WYNNE, Peter M., 1997. *Management information systems and performance measurement for the electronic library: eLib supporting study, MIEL 2, final report* [online]. Preston. Available from: <http://www.ukoln.ac.uk/dlis/models/studies/mis/mis.rtf>.

BROPHY, Peter and WYNNE, Peter M., 1999. Performance indicators for the electronic library. In: *SCONUL newsletter*. 1999. Vol. 16, no. Spring, pp. 3–5.

BUSH, Vannevar, 1945. As We May Think. In: *The Atlantic Monthly* [online]. 1945. no. 1 July. Available from: <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>.

CAMPBELL, Nicole and U.S., Library and Information Technology Association, 2001. *Usability assessment of library-related Web sites: methods and case studies* [online]. S.l.: s.n. [Accessed 2 November 2013]. ISBN 0838981577. Available from: http://books.google.pt/books/about/Usability_assessment_of_library_related.html?id=DdsmAQAAMAAJ&pgis=1.

CASE, Donald, 2012. *Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior (Library and Information Science)* [online]. S.l.: Emerald Group Publishing. [Accessed 23 December 2013]. ISBN 1780526547. Available from: <http://www.amazon.com/Looking-Information-Research-Seeking-Behavior/dp/1780526547>.

CHOUDHURY, SAYEED HOBBS, Benjamin and LORIE, Mark, 2002. A Framework for Evaluating Digital Library Services. In: *D-Lib Magazine* [online]. 2002. Vol. 8, no. 7/8. Available from: <http://www.dlib.org/dlib/july02/choudhury/07choudhury.html>.

CHOWDHURY, G.G.; CHOWDHURY, Sudatta – *Information sources and searching on the world wide web*. London : Library Association Publishing, 2001. XVIII, 174 p. ISBN 1-85604-394-0.

CHOWDHURY, G.G.; CHOWDHURY, Sudatta, 2003. *Introduction to digital libraries*. S.l.: s.n. ISBN 1-85604-465-3.

CHOWDHURY, G.G. and CHOWDHURY, Sudatta, 2011. *Information users and usability in the digital age*. 1st. London: Facet Publishing. ISBN 978-1-85604-597-1.

CHOWDHURY, Gobinda G., 2004. Access and Usability Issues of Scholarly Electronic Publications. In: GORMAN, G. E. and ROWLAND, F. (eds.), *Scholarly publishing in an electronic era: international yearbook of Library and Information Management* [online]. S.l.: Facet Publishing. pp.77–98. Available from: http://www.cis.strath.ac.uk/cis/research/publications/papers/strath_cis_publication_315.pdf.

CHOWDHURY, Gobinda, KOO, Chris and HUNTER, Jane (eds.), 2010. *The Role of Digital Libraries in a Time of Global Change* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg. [Accessed 21 November 2013]. ISBN 978-3-642-13653-5. Available from: <http://www.springerlink.com/index/10.1007/978-3-642-13654-2>.

CHOWDHURY, Gobinda, MCMENEMY, David and POULTER, Alan, 2008. MEDLIS: Model for Evaluation of Digital Libraries and Information Services. In: *World Digital Libraries* [online]. 2008. Vol. 1, no. 1. Available from: <http://iospress.metapress.com/content/y58640h474v483w7/>.

CHOWDHURY, Sudatta, LANDONI, Monica and GIBB, Forbes, 2006. Usability and impact of digital libraries: a review. In: *Online Information Review* [online]. 11 January 2006. Vol. 30, no. 6, pp. 656–680. [Accessed 5 September 2013]. DOI 10.1108/14684520610716153. Available from:

<http://www.emeraldinsight.com/journals.htm?issn=1468-4527&volume=30&issue=6&articleid=1583935&show=html>.

CHOWDHURY, G., MCMENEMY, D. AND POULTER, A., 2008. MEDLIS: Model for Evaluation of Digital Libraries and Information Services. *World Digital Libraries*, 1(1).

CHU, Heting, 2003. *Information Representation and Retrieval in the Digital Age* [online]. 1st. S.l.: Information Today for the American Society for Information and Technology. [Accessed 24 April 2013]. ISBN 1-57387-172-9. Available from: <http://www.emeraldinsight.com/journals.htm?issn=1468-4527&volume=28&issue=5&articleid=1499426&show=html>.

COVEY, Denise Troll, 2002. *Usage and usability assessment : library practices and concerns* [online]. 1st ed. Washington D.C.: Council on Library and Information Resources. ISBN 1-887334-89-0. Available from: <http://www.clir.org/pubs/reports/pub105/contents.html>.

CULLEN, Rowena J. and CALVERT, Philip J., 1995. Stakeholder perceptions of university library effectiveness. In: *The Journal of Academic Librarianship*. 1995. Vol. 21, no. 6, pp. 438–448. DOI 10.1016/0099-1333(95)90087-X.

CURRÁS, Emilia, 2010. *Ontologias, taxonomia e tesouros em teoria de sistemas e sistemática*. Brasília: Thesaurus Editora. ISBN 978-85-7062-968-5.

DILLON, A., 1999. Evaluating on TIME: A framework for the expert evaluation of digital library interface usability. In: *International Journal on Digital Libraries*. 1999. Vol. 2, no. 2/3, pp. 170–177.

DOBREVA, Milena and CHOWDHURY, Sudatta, 2010. A user-centric evaluation of the Europeana digital library. In: [online]. S.l.: Springer-Verlag. 21 June 2010.

pp. 148–157. [Accessed 21 November 2013]. Available from:
<http://dl.acm.org/citation.cfm?id=1875689.1875714>.

DUMAS, Joseph F. and REDISH, Janice C., 1993. *A Practical Guide to Usability Testing* [online]. S.l.: Greenwood Publishing Group Inc. [Accessed 8 November 2013]. ISBN 0893919918. Available from:
<http://dl.acm.org/citation.cfm?id=562615>.

ESCOLAR SOBRINHO, Hipólito, 1990. *Historia de las bibliotecas*. Madrid: Ediciones Pirámide. ISBN 84-368-0535-6.

FERNANDES, Maria Eugénia Matos and RIBEIRO, Lígia Maria, 2009. *O Repositório Aberto da Universidade do Porto* [online]. 2009. S.l.: s.n. [Accessed 20 November 2013]. Available from: <http://repositorio-aberto.up.pt/handle/10216/18147>.

FERRO, Nicola and SILVELLO, Gianmaria, 2013. NESTOR: A formal model for digital archives. In: *Information Processing & Management* [online]. November 2013. Vol. 49, no. 6, pp. 1206–1240. [Accessed 5 September 2013]. DOI 10.1016/j.ipm.2013.05.001. Available from:
<http://dx.doi.org/10.1016/j.ipm.2013.05.001>.

FIELDS, Bob, KEITH, Suzette and BLANFORD, Ann, 2003. *Usability evaluation of digital libraries : a tutorial*. Middlesex.

FOX, Edward A., 2005. Digital Libraries Initiative (DLI) Projects 1994-1999. In: *Bulletin of the American Society for Information Science and Technology* [online]. 31 January 2005. Vol. 26, no. 1, pp. 7–11. [Accessed 11 December 2013]. DOI 10.1002/bult.135. Available from: <http://doi.wiley.com/10.1002/bult.135>.

FUHR, Norbert, TSAKONAS, Giannis, AALBERG, Trond, AGOSTI, Maristella, HANSEN, Preben, KAPIDAKIS, Sarantos, KLAS, Claus-Peter, KOVÁCS, László, LANDONI, Monica, MICSIK, András, PAPTAEODOROU, Christos, PETERS, Carol and SØLVBERG, Ingeborg, 2007. Evaluation of digital libraries. In: *Int J Digit Libr* [online]. 2007. Vol. 8, pp. 21–38. DOI 10.1007/s00799-007-0011-z. Available from: <http://comminfo.rutgers.edu/~tefko/Courses/Zadar/Readings/Fuhr Eval of dig lib Int J on DL 2007.pdf>.

FURTADO, Elizabeth, et al., 2003. Improving usability of an online learning system by means of multimedia, collaboration, and adaptation resources. In: *Usability Evaluation of Online Learning Programs* [online]. S.l.: Information Science Publ. pp. 69–86. [Accessed 10 November 2013]. ISBN 1591401054. Available from: <http://www.google.pt/books?hl=pt-PT&lr=&id=TElc5gdmYJEC&pgis=1>.

GARROD, Penny and KINNELL, Margaret, 1997. Benchmarking development needs in the LIS sector. In: *Journal of Information Science* [online]. 1997. Vol. 23, no. 2, pp. 111–118. Available from: <http://jis.sagepub.com/content/23/2/111.abstract>.
GLUCK, Myke, 1997. A Descriptive Study of the Usability of Geospatial Metadata. In: *Annual Review of OCLC Research* [online]. 1997. Available from: <http://library.oclc.org/cdm/ref/collection/p267701coll27/id/70>.

GLUCK, M., 1997. A Descriptive Study of the Usability of Geospatial Metadata. *Annual Review of OCLC Research*.

HARRIS, Michael H., 1995. *History of Libraries in the Western World*. London: The Scarecrow Press. ISBN 0-8108-2972-X.

HARTSON, H. Rex, SHIVAKUMAR, Priya and PÉREZ-QUIÑONES, Manuel A., 2004. Usability inspection of digital libraries: a case study. In: *International Journal on Digital Libraries* [online]. 23 July 2004. Vol. 4, no. 2, pp. 108–123. [Accessed 4 December 2013]. DOI 10.1007/s00799-003-0074-4. Available from: <http://link.springer.com/10.1007/s00799-003-0074-4>.

HEARST, Martí A. – User interfaces and visualization. In BAEZA-YATES, Ricardo ; RIBEIRO-NETO, Berthier – **Modern information retrieval**. 1st ed. Harlow : ACM Press, 1999. p. 257-323. ISBN 0-201-39829-X.

HIX, Deborah and HARTSON, H. Rex, 1993. Developing user interfaces: ensuring usability through product & process. In: [online]. 21 May 1993. [Accessed 11 November 2013]. Available from: <http://dl.acm.org/citation.cfm?id=126701>.

HJØRLAND, Birger, 2007. Semantics and knowledge organization. In: *Annual Review of Information Science and Technology* [online]. 2007. Vol. 41, pp. 367–405. Available from: https://www.academia.edu/1912560/Semantics_and_knowledge_organization.

INGWERSEN, Peter, 1996. Cognitive perspectives of information retrieval interaction: elements of a cognitive theory. In: *Journal of Documentation* [online]. 31 December 1996. Vol. 52, no. 1, pp. 3–50. [Accessed 22 December 2013]. DOI 10.1108/eb026960. Available from: <http://www.emeraldinsight.com/journals.htm?issn=0022-0418&volume=52&issue=1&articleid=1650202&show=html>.

ISIDRO, Rui Orlando, FERNANDES, Francisco José, BEÇA, Pedro Miguel, PINTO, Joaquim Sousa, SILVA, Lídia Oliveira and DUARTE, A. Manuel de Oliveira, 2012. *ABC: Arquivo Bibliográfico para publicações Científicas* [online]. 12 April 2012. S.l.: s.n. [Accessed 20 November 2013]. Available from: <http://revistas.ua.pt/index.php/revdeti/article/view/1593>. ~

ISO, 2003a. *ISO 2789. 2003, Information and documentation – International library statistics*. 2003. Genève: ISO.

ISO, 2003b. *ISO/TR 20983. 2003, Information and documentation - Performance indicators for electronic library services*. 2003. Genève: ISO.

ISO 11620, 1998. *Library performance indicators*. 1998. Genève: s.n.

ISO 9241-11. 1998 - Ergonomic requirements for office work with visual display terminals. Part 11: Guidance on usability. Genève : ISO.

IVORY, Melody Y. ; HEARST, Marti – The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*. New York : ACM Press. ISSN 0360-0300. 33:4 (2001) 470-516.

JANSEN, Bernard Jim – *Understanding user-web interactions via web analytics*. [s.l.] : Morgan & Claypool Publishers, 2009. XIII,101 p. ISBN 9781598298512.

JÄVERLIN, Kalervo and INGWERSEN, Peter, 2004. Information Seeking Research Needs Extension towards Tasks and Technology. In: *Information Research* [online]. 2004. Vol. 10, no. 1. [Accessed 20 December 2013]. Available from: <http://www.informationr.net/ir/10-1/paper212.html>.

JENG, Judy, 2005. What is usability in the context of the digital library and how can it be measured? In: *Information Technology and Libraries* [online]. 2005. Vol. 24, no. 2, pp. 47–56. Available from: http://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1043&context=slis_pub.

JONES, A, SCANLON, E, TOSUNOGLU, C, MORRIS, E, ROSS, S, BUTCHER, P and GREENBERG, J, 1999. Contexts for evaluating educational software. In: *Interacting with Computers* [online]. 1999. Vol. 11, no. 5, pp. 499–516. [Accessed 6 December 2013]. Available from: <http://www.sciencedirect.com/science/article/pii/S0953543898000642>.

KANTNER, Laurie and ROSENBAUM, Stephanie, 1997. Usability studies of WWW sites. In: *Proceedings of the 15th annual international conference on Computer documentation - SIGDOC '97* [online]. New York, New York, USA: ACM Press. 1 October 1997. pp. 153–160. [Accessed 2 November 2013]. Available from: <http://dl.acm.org/citation.cfm?id=263367.263388>.

KEITH, Suzette, BLANDFORD, Ann, FIELDS, Bob and THENG, Yin Leng, 2002. No. In: *Ann Blandford & George Buchanan (Eds.), JCDL'02 Workshop on Usability of Digital Libraries* [online]. S.l.: s.n. 2002. Available from: <http://www.ucl.ac.uk/annb/docs/Keith15.pdf>.

KELLY, Brian – Implementing a quality assurance methodology for digital library programmes. *The journal of information and knowledge management systems*. ISSN 0305-5728. 35:1/2 (2005) 97-104.

KENGERI, Rekha, SEALS, Cheryl D., HARLEY, Hope D., REDDY, Himabindu P. and FOX, Edward A., 1999. Usability study of digital libraries: ACM, IEEE-CS, NCSTRL, NDLTD. In: *International Journal on Digital Libraries* [online]. 1 September 1999. Vol. 2, no. 2-3, pp. 157–169. [Accessed 11 November 2013]. DOI 10.1007/s007990050044. Available from: <http://link.springer.com/10.1007/s007990050044>.

KIM, K., 2002. A Model-based Approach to Usability Evaluation for Digital Libraries. In: *Ann Blandford & George Buchanan (Eds.), JCDL'02 Workshop on Usability of Digital Libraries* [online]. 2002. Available from: <http://www.ucl.ac.uk/annb/docs/Kim33.pdf>.

KLAS, C.P. [et al.] - A logging scheme for comparative digital library evaluation. *Research and Advanced Technology for Digital Libraries*. ISSN 0302-9743. 4172 (2006) 267-278.

KUROSU, Masaaki and KASHIMURA, Kaori, 1995. Apparent usability vs. inherent usability. In: *Conference companion on Human factors in computing systems - CHI '95* [online]. New York, New York, USA: ACM Press. 7 May 1995. pp. 292–293. [Accessed 29 November 2013]. Available from: <http://dl.acm.org/citation.cfm?id=223355.223680>.

LANCASTER, F.W. - *Indexing & abstracting in theory & practice*. 3rd ed. Urbana-Champaign : University of Illinois, 2003. ISBN 0878451226.

LESK, Michael, 1997. *Practical Digital Libraries: Books, Bytes and Bucks* [online]. 1st ed. S.l.: Morgan Kaufmann Publishers In. [Accessed 11 December 2013]. ISBN 1558604596. Available from: <http://www.amazon.co.uk/Practical-Digital-Libraries-Multimedia-Information/dp/1558604596>.

LESK, Michael, 2012. A personal history of digital libraries. In: *Library Hi Tech* [online]. 16 November 2012. Vol. 30, no. 4, pp. 592–603. [Accessed 21 April 2013]. DOI 10.1108/07378831211285077. Available from: <http://www.emeraldinsight.com/journals.htm?issn=0737-8831&volume=30&issue=4&articleid=17065298&show=html>.

LEVI, Michael D. and CONRAD, Frederick G., 1997. Usability testing of world wide web sites. In: *CHI '97 extended abstracts on Human factors in computing systems looking to the future - CHI '97* [online]. New York, New York, USA: ACM Press. 22 March 1997. pp. 227. [Accessed 2 November 2013]. Available from: <http://dl.acm.org/citation.cfm?id=1120212.1120358>.

LØRRING, Leif and KAJBERG, Leif (eds.), 2005. *European Curriculum Reflections on Library and Information Science Education*. Copenhagen: The Royal School of Library and Information Science. ISBN 87-7415-292-0.

MADLE, Gemma, KOSTKOVA, Patty, MANI-SAADA, Jane and WEINBERG, Julius, 2003. Development of a methodology to evaluate the impact of a medical digital library on user knowledge, attitude and behaviour. In: *Mednet 2003, 4-7 December* [online]. S.l.: s.n. 2003. Available from: <http://www.hon.ch/Mednet2003/abstracts/491753869.html>.

MALCLÉS, Louise-Noëlle, 1985. *Manuel de bibliographie*. 4me. ed. rev. et augm. Paris: Presses Universitaires de France.

MANNING, Christopher D.; RAGHAVAN, Prabhakar; SCHÜTZE, Hinrich – *Introduction to information retrieval*. Cambridge University Press, 2008. ISBN 978-0-521-86571-5.

MARCHIONINI, Gary, 2000. Evaluating Digital Libraries: A Longitudinal and Multifaceted View. In: *Library Trends* [online]. 2000. Vol. 49, no. 2, pp. 304–333. [Accessed 5 December 2013]. Available from: <https://www.ideals.illinois.edu/handle/2142/8336>.

MARCHIONINI, Gary, 1995. *Information seeking in electronic environments*. Cambridge: Cambridge University Press.

MARCHIONINI, Gary and KOMLODI, Anita, 1997. Design of Interfaces for Information Seeking. In: *Annual Review of Information Science and Technology (ARIST)* [online]. 30 November 1997. Vol. 33, pp. 89–130. [Accessed 31 October 2013]. Available from: <http://eric.ed.gov/?id=EJ610153>.
citado por Chowdhury - cf p. 154

MARSHALL, J G, 1992. The impact of the hospital library on clinical decision making: the Rochester study. In: *Bull Med Libr Assoc*. 1992. Vol. 80, no. 2, pp. 169–178.

MELUCCI, M. AND BAEZA-YATES, R., 2011. *Advanced Topics in Information Retrieval* 1st ed., Heidelberg: Springer.

MOOERS, C.N., 1951. Zatocoding applied to mechanical organization of knowledge. *American Documentation*, 2, pp.20–32.

MOORE, Nick, 1989. *Comment mesurer l'efficacité des bibliothèques publiques*. Paris: UNESCO.

PGI-89/WS/3

NIELSEN, Jakob, 1993. *Usability Engineering* [online]. S.l.: Academic Press. [Accessed 11 November 2013]. ISBN 0080520294. Available from: <http://www.google.pt/books?hl=pt-PT&lr=&id=DBOowF7LqIQc&pgis=1>.

NIELSEN, Jakob and MACK, Robert L., 1994. *Usability Inspection Methods* [online]. S.l.: Wiley. [Accessed 2 November 2013]. ISBN 0471018775. Available from: <http://www.amazon.com/Usability-Inspection-Methods-Jakob-Nielsen/dp/0471018775>.

NIELSEN, Jakob and PERNICE, Kara, 2009. *Eyetracking Web Usability* [online]. 1st. Thousand Oaks, CA, USA: New Riders Publishing. [Accessed 8 December 2012]. ISBN 0321498364, 9780321498366. Available from: <http://dl.acm.org/citation.cfm?id=1823564>.

NORLIN, Elaina, 2000. Reference Evaluation: A Three-Step Approach--Surveys, Unobtrusive Observations, and Focus Groups. In: *Coll. res. libr.* [online]. 1 November 2000. Vol. 61, no. 6, pp. 546–553. [Accessed 31 October 2013]. Available from: <http://crl.acrl.org/content/61/6/546.short>.

OLIVEIRA, Miguel, PEREIRA, Pedro Miguel dos Santos Beça, PINTO, Joaquim Sousa and DUARTE, Aníbal Manuel de Oliveira, 2004. eABC: scientific publications bibliographic archives. In: ENGELEN, Jan, COSTA, Sely M. S. and 220

MOREIRA, Ana Cristina S. (eds.), *Building Digital Bridges: Linking Cultures, Commerce and Science: Proceedings of the 8th ICC/IFIP International Conference on Electronic Publishing held in Brasília* [online]. Brasília: ELPUB 2004. 2004. [Accessed 20 November 2013]. Available from: <http://ria.ua.pt/handle/10773/6559>.

OULANOV, Alexei and PAJARILLO, Edmund J.Y., 2002. CUNY+ Web: usability study of the Web-based GUI version of the bibliographic database of the City University of New York (CUNY). In: *The Electronic Library* [online]. 12 January 2002. Vol. 20, no. 6, pp. 481–487. [Accessed 11 November 2013]. DOI 10.1108/02640470210454001. Available from: <http://www.emeraldinsight.com/journals.htm?issn=0264-0473&volume=20&issue=6&articleid=861982&show=html>.

PAN, B., GAY, G., SAYLOR, J., HEMBROOKE, H. and HENDERSON, D., 2004. Usability, learning, and subjective experience: user evaluation of K-MODDL in an undergraduate class. In: *Proceedings of the 2004 Joint ACM/IEEE Conference on Digital Libraries*. S.l.: s.n. 2004. pp. 188–189.

PATTON, Michael Quinn, 1990. *Qualitative Evaluation and Research Methods* [online]. S.l.: SAGE Publications. [Accessed 20 December 2013]. ISBN 0803937792. Available from: http://books.google.pt/books/about/Qualitative_Evaluation_and_Research_Meth.html?id=XRO2AAAAIAAJ&pgis=1.

PEARROW, Mark, 2006. *Web Site Usability Handbook (Internet Series)* [online]. S.l.: Cengage Learning. [Accessed 2 November 2013]. ISBN 1584504692. Available from: <http://www.amazon.com/Site-Usability-Handbook-Internet-Series/dp/1584504692>.

PERNICE, Kara and NIELSEN, Jakob, 2009. *Eyetracking methodology* [online]. Fremont: Nielsen Norman Group. Available from:

http://media.nngroup.com/media/reports/free/How_to_Conduct_Eyetracking_Studies.pdf

POLL, Roswitha ; BOEKHORST, Peter te - *Measuring quality : international guidelines for performance measurement in academic libraries*. [1ª ed.]. München : K G Saur, 1996. 171 p. (IFLA Publications ; 76). ISBN 3-598-21800-1.

POPP, Mary Pagliero, 2001. *Testing Library Web Sites : ARL Libraries Weigh In* [online]. 2001. S.l.: s.n. Available from: <http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/pdf/popp.pdf>. Paper presented at the Association of College and Research Libraries, 10th National Conference, Denver, Colorado, 15-18 March 2001

RIBEIRO, Cristina, LOPES, João Correia, SILVA, João Rocha da and FERNANDES, Maria Eugénia Matos, 2012. *Repositório de dados na U.Porto: um fluxo de curadoria suportado numa extensão do DSPACE* [online]. 2012. S.l.: s.n. [Accessed 20 November 2013]. ISBN <http://hdl.handle.net/10216/65088>. Available from: <http://repositorio-aberto.up.pt/handle/10216/65088>.

RIBEIRO, Cristina, RODRIGUES, Eloy, FERNANDES, Maria Eugénia Matos and SARAIVA, Ricardo, 2010. *Os Repositórios de Dados Científicos: Estado da Arte* [online]. 2010. S.l.: s.n. [Accessed 20 November 2013]. ISBN <http://hdl.handle.net/10216/23806>. Available from: <http://repositorio-aberto.up.pt/handle/10216/23806>.

RIBEIRO, Fernanda; SILVA, Armando Malheiro da – A avaliação de informação : uma operação metodológica. *Páginas a&b : arquivos e bibliotecas*. Lisboa : Gabinete de Estudos a&b. ISSN 0873-5670. 14 (2004) 7-37.

RODRIGUES, Eloy, SWAN, Alma and BAPTISTA, Ana Alice, 2013. *Uma Década de Acesso Aberto na UMinho e no Mundo* [online]. S.l.: Universidade do Minho,

Serviços de Documentação. [Accessed 20 November 2013]. ISBN 978-989-98704-1-3. Available from: <http://repositorium.sdum.uminho.pt/handle/1822/26144>.

ROSSON, Mary Beth and CARROLL, John M., 2001. *Usability Engineering: Scenario-Based Development of Human-Computer Interaction (Interactive Technologies)* [online]. S.l.: Morgan Kaufmann. [Accessed 2 November 2013]. ISBN 1558607129. Available from: <http://www.amazon.com/Usability-Engineering-Scenario-Based-Human-Computer-Technologies/dp/1558607129>.

RUBIN, Jeffrey, 1994. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests* [online]. 1st. S.l.: Wiley. ISBN 9780471594031. Available from:
http://www.amazon.com/exec/obidos/ASIN/0471594032/qid=1045724394/sr=2-1/ref=sr_2_1/104-1975213-1712725.

SANDUSKY, Robert J., 2002. Digital Library Attributes: Framing Usability Research. In: *Ann Blandford & George Buchanan (Eds.), JCDL'02 Workshop on Usability of Digital Libraries* [online]. S.l.: s.n. 2002. Available from: <http://www.ucl.ac.uk/annb/docs/Sandusky35.pdf>.

SANTOS, Susete Margarida de Jesus Lopes Serra dos, 2011. *A biblioteca digital como recurso informacional : uma análise da sua aplicabilidade ao apoio ao ensino-aprendizagem e à investigação na Universidade de Aveiro* [online]. S.l.: s.n. [Accessed 20 November 2013]. Available from: <https://estudogeral.sib.uc.pt/jspui/handle/10316/19049>.

Dissertação de mestrado em Informação, Comunicação e Novos Media, apresentada à Faculdade de Letras da Universidade de Coimbra

SARACEVIC, Tefko, 1996. Ciência da informação: origem, evolução e relações. In: *Perspectivas em Ciência da Informação* [online]. 3 June 1996. Vol. 1, no. 1, pp. 41–

62. [Accessed 19 December 2013]. Available from:
<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235>.

SARACEVIC, Tefko - Digital Library Evaluation : Toward an Evolution of Concepts. *Library Trends*. ISSN 0024-2594. 49:2 (2000) 350-369.

SERRANO, Anabela – A avaliação de serviços de informação do ensino superior. *Páginas a&b : arquivos e bibliotecas*. Lisboa : Gabinete de Estudos a&b. ISSN 0873-5670. 12 (2003) 7-65.

SERRANO, Anabela, 2011. Information Retrieval Educational Goals in Library and Information Science and in Health Sciences. In: *Information Services and Use* [online]. 2011. Vol. 31, no. 3/4, pp. 131–138. Available from:
<http://iospress.metapress.com/content/w74430xg44r2mx51/?p=517390da9cf641e98320903070b2f2fc&pi=5>.

SHACKEL, B, 1986. Ergonomics in design for usability. In: *Proceedings of the Second Conference of the British Computer Society, human computer interaction specialist group on People and computers: designing for usability* [online]. S.l.: Cambridge University Press. 1 December 1986. pp. 44–64. [Accessed 11 November 2013]. Available from:
<http://dl.acm.org/citation.cfm?id=17324.24077>.

SMITH, Alastair G. – Search features of digital libraries. *Information Research*. ISSN 1368-1613. 5:3 (2000). Acessível na Internet: <http://informationr.net/ir/5-3/paper73.html>

SHNEIDERMAN, Ben ; PLAISANT, Catherine – *Designing the user interface : strategies for effective human-computer interaction*. Boston : Pearson, cop. 2005. XVIII, 652 p. ISBN 0-321-26978-0.

SILVA, A.M. DA AND RIBEIRO, F., 2002. *Das “ciências” documentais à ciência da informação: ensaio epistemológico para um novo modelo curricular*, Edições Afrontamento.

SNYDER, Carolyn, 2003. *Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces (Interactive Technologies)* [online]. S.l.: Morgan Kaufmann. [Accessed 2 November 2013]. ISBN 1558608702. Available from: <http://www.amazon.com/Paper-Prototyping-Interfaces-Interactive-Technologies/dp/1558608702>.

SUMNER, Tamara and DAWES, Melissa, 2001. Looking at digital library usability from a reuse perspective. In: *Proceedings of the first ACM/IEEE-CS joint conference on Digital libraries - JCDL '01* [online]. New York, New York, USA: ACM Press. 1 January 2001. pp. 416–425. [Accessed 4 December 2013]. Available from: <http://dl.acm.org/citation.cfm?id=379437.379742>.

SUMNER, Tamara, KHOO, Michael, RECKER, Mimi and MARLINO, Mary, 2003. Understanding educator perceptions of quality in digital libraries. In: [online]. S.l.: IEEE Computer Society. 27 May 2003. pp. 269–279. [Accessed 1 December 2013]. Available from: <http://dl.acm.org/citation.cfm?id=827140.827188>.

SVENONIUS, Elaine, [no date]. The Epistemological Foundations of Knowledge Representations. In: *Library Trends* [online]. Vol. 52, no. 3, pp. 571–587. [Accessed 19 December 2013]. Available from: <https://www.ideals.illinois.edu/handle/2142/1691>.

THENG, Yin Leng, MOHD-NASIR, Norliza and THIMBLEBY, Harold, 2000. Purpose and usability of digital libraries. In: *Proceedings of the fifth ACM conference on Digital libraries - DL '00* [online]. New York, New York, USA: ACM Press. 1 June 2000. pp. 238–239. [Accessed 30 November 2013]. Available from: <http://dl.acm.org/citation.cfm?id=336597.336674>.

THENG, Yin-Leng, CHAN, Mei-Yee, KHOO, Ai-Ling and BUDDHARAJU, Raju, 2004. *Design and Usability of Digital Libraries* [online]. S.l.: IGI Global. [Accessed 4 December 2013]. ISBN 9781591404415. Available from: <http://www.igi-global.com/chapter/quantitative-qualitative-evaluations-singapore-national/8147>.

THOMAS, R. Murray, 2003. *Blending Qualitative and Quantitative Research Methods in Theses and Dissertations* [online]. Thousand Oaks, CA, USA: Corwin Press. [Accessed 20 December 2013]. ISBN 0761939326. Available from: http://books.google.pt/books/about/Blending_Qualitative_and_Quantitative_Re.html?id=qx-yzSILfbMC&pgis=1.

TSAKONAS, G.; PAPTAEODOROU, C. - Analysing and evaluating usefulness and usability in electronic information services. *Journal of Information Science*. ISSN 0165-5515. 32:5 (2006) 400-419.

VAN HOEK, Wilko and MAYR, Philipp, 2013. Assessing Visualization Techniques for the Search Process in Digital Libraries. In: STEFAN ANDREAS KELLER, SCHNEIDER, René and VOLK, Benno (eds.), *Wissensorganisation mit digitalen Technologien* [online]. S.l.: DeGruyter. Available from: <http://arxiv.org/abs/1304.4119>.

VAN RIJSBERGEN, C. J., 1979. *Information retrieval* [online]. S.l.: Butterworths. [Accessed 20 December 2013]. ISBN 0408709294. Available from: http://books.google.pt/books/about/Information_retrieval.html?id=t-pTAAAAMAAJ&pgis=1.

WILSON, T.D., 1999. Models in information behaviour research. *Journal of Documentation*, 55(3), pp.249–270.

WINKWORTH, Ian - Performance indicators for electronic services : an outline of SCONUL and UCISA work. *SCONUL newsletter*. ISSN 13530429. 16 (Spring 1999) 5.

XIE, Hong Iris - Evaluation of digital libraries : Criteria and problems from users' perspectives. *Library & Information Science Research*. ISSN 0740-8188. 28:3 (2006) 433-452.

YOUNG, Peter R. - Measurement of electronic services in libraries : statistics for the digital age. *IFLA journal*. ISSN 03400352. 24:3 (1998) 157-160.

ZHANG, Tao, MARON, Deborah J. and CHARLES, Christopher C., 2013. Usability evaluation of a research repository and collaboration website. In: *Journal of Web Librarianship* [online]. 2013. Vol. 7, no. 1, pp. 58–82. DOI 10.1080/19322909.2013.739041. Available from: http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1061&context=lib_fsdocs.

ZHAO, Yumin, NIU, Zhendong, CAO, Yujuan and DAI, Lin, 2010. Research on Evaluation of Digital Library. In: *2010 International Conference on Data Storage and Data Engineering* [online]. S.l.: IEEE. 9 February 2010. pp. 123–126. [Accessed 29 November 2013]. Available from: <http://www.computer.org/csdl/proceedings/dsde/2010/3958/00/3958a123-abs.html>.

Anexos

Anexo 1

Tabela 1 - Projetos e programas de investigação em avaliação de serviços de informação²³⁹

Identificação do Projeto	Datas	Origem	Designação e Descrição
CAVAL	(1995-1998)	Projeto australiano	
CAUL	(1995)	Projeto australiano	
CAMILE ²⁴⁰	(1998-1999)	Projeto europeu	(Concerted Action on Management Information for Libraries in Europe) Ação concertada do Programa de Aplicações Telemáticas da Comissão Europeia. Teve a seu cargo a difusão e promoção dos resultados de quatro projetos europeus: DECIDE, EQLIPSE, DECIMAL, e MINSTREL. O projeto EQUINOX desenvolveu-se tendo por base as conclusões destes projetos.
CNI Projects ²⁴¹		Projetos americanos	
Cranfield project	(1997-1998)	Projeto inglês	(Project on performance indicators for academic libraries)

²³⁹ Uma versão anterior desta tabela foi publicada num artigo de publicação periódica em 2003 pela autora da tese - cf. Serrano (2003).

²⁴⁰ Acessível na Internet: <<http://www.staff.dmu.ac.uk/~camile/>>.

²⁴¹ Acessível na Internet: <<http://www.cni.org/projects/assessing/reports>>.

DECIDE	(1995-1997)	Projeto europeu	(Decision Support Models : a Decision Support System for European Academic and Public Libraries)
DELOS	(2004-2007)	Projeto europeu	Inclui 7 “clusters”: Digital Library Architecture; Information Access and Personalization; Audio/Visual and Non-traditional Objects; User interfaces and Visualization; Knowledge Extraction and Semantic Interoperability; Preservation; Evaluation
DECIMAL	(1995-1997)	Projeto europeu	(Decision-making in Libraries : decision research for the development of integrated library systems)
DETYA/EIP		Projeto australiano	(‘Best practice in Australian university libraries’ project)
DigiQUAL	(2003-)	Projeto americano	Projeto parceiro do LibQUAL+ dedicado à avaliação de bibliotecas digitais
DL.org	(2008-)	Projeto europeu	Projeto que continuou o projeto DELOS. Inclui 6 grupos de trabalho: Content; Functionality; Users; Policy; Quality; Architecture.
E-Metrics	(2000-2001)	Projeto americano	Projeto desenvolvido pela Association of Research Libraries (ARL).
EQLIPSE	(1995-1997)	Projeto europeu	(Evaluation and Quality in Library Performance: System for Europe) Projeto coordenado pela University of Central Lancashire
EQUINOX ²⁴² -	(1998-2000)	Projeto europeu	(Library Performance Measurement and Quality Management System)
LibQUAL+Project	(1999-)	Projeto	Projeto desenvolvido pela

²⁴² Acessível na Internet: <<http://equinox.dcu.ie/>>.

		americano	Association of Research Libraries (ARL) a partir do projeto SERVQUAL
MIEL ²⁴³	(1998)	Projeto inglês	(Management Information for the Electronic Library)
MINSTREL	(1995-1997)	Projeto europeu	
NACUBO Benchmark Program	(1998-1999)	Projeto americano	Programa levado a cabo pela “National Association of College and University Business Officers”
PROLIB – PI	(1994-1995)	Projeto europeu	
Quality	(dentro do período de 1993-1998)	Projeto europeu	Gerido pelo Centre for Research in Library and Information Management (CERLIM) na época em que estava sediado na University of Central Lancashire
SCONUL Benchmarking Pilot Project	(1997-1999)	Projeto inglês	Cranfield University
SCONUL User Satisfaction Project	(1998-1999)	Projeto inglês	Cranfield University
SERVQUAL	(1983-2000)	Projeto americano	Projeto desenvolvido pela Association of Research Libraries (ARL)

²⁴³ Acessível na Internet: <<http://www.ukoln.ac.uk/dlis/models/studies/mis/mis.rtf>>.

Anexo 2

Tabela 2 - Critérios de avaliação da usabilidade por autor(es) e por ano

Critério de avaliação da usabilidade ou atributo de usabilidade	Autores	Ano
Adaptabilidade	Oulanov e Pajarillo	2002
Afeto ou <i>affect</i>	Oulanov e Pajarillo	2002
Ajuda e recuperação de erros	Blanford e Buchanan	2003
Apreensibilidade	Shackel	1986
	Booth	1989
	Nielsen	1993
	Hix e Hartson	1993
	Kengeri, et al	1999
	Blanford e Buchanan	2003
Aprendizagem	Furtado, et al	2003
Atingir objetivos	Blanford e Buchanan	2003
Atitude	Booth	1989
Atitude do utilizador	Shackel	1986
Capacidade de retenção ou <i>retainability</i>	Hix e Hartson	1993
Capacidade de ser utilizável ou <i>Useableness</i>	Gluck	1997
Contexto	Blanford e Buchanan	2003
Controlo	Oulanov e Pajarillo	2002
Desempenho a longo prazo	Hix e Hartson	1993

Eficácia	Shackel	1986
	Booth ²⁴⁴ - igual a <i>Ease of use</i>	1989
	ISO	
	Kengeri, et al	1998
		1999
Eficácia da interface	Kim ²⁴⁵	2002
Eficiência	Nielsen	1993
	ISO	1998
	Oulanov e Pajarillo	2002
Eficiente de usar	Brinck, et al	2002
Erros	Nielsen	1993
Experiência do utilizador	Blanford e Buchanan	2003
Fácil de aprender	Brinck, et al	2002
Fácil de recordar	Brinck, et al	2002
Fácil recuperação de erros	Nielsen	1993
Facilidade de uso	Furtado, et al	2003
Flexibilidade	Shackel	1986
Funcionalmente correto	Brinck, et al	2002
Initial performance	Hix e Hartson	1993
Likeability ou apreciabilidade	Kengeri, et al	1999
Memorabilidade	Nielsen	1993
Primeira impressão	Hix e Hartson	1993
Satisfação	Nielsen	1993
	ISO	1998
Satisfação do utilizador a longo prazo	Hix e Hartson	1993
Subjetivamente agradável	Brinck, et al	2002
Taxa de erro baixa	Nielsen	1993
Tolerante a erros	Brinck, et al	2002

²⁴⁴ Citado por Jeng (2005).

²⁴⁵ Kim, Kyunghye. 2002. A model of digital library information seeking process (DLISP model) as a frame for classifying usability problems. PhD diss., Rutgers Univ. Citada por Jeng (2005).

<i>User lostness</i>	Nielsen	1993
Utilidade	Booth	1989
	Gluck	1997
	Kengeri, et al	1999
	Oulanov e Pajarillo	2002
Utilização de funcionalidades avançadas	Hix e Hartson	1993

Anexo 3

Tabela 3 – Repositórios digitais sediados em Portugal em 2009 com registo no ROAR e no OpenDOAR²⁴⁶

Nome do Repositório Digital	Organização Responsável	Ano de registo no ROAR	ID no ROAR	ID de Registo no OpenDOAR	Software	Número de documentos depositados (2009)
ARCA – IGC Repository	Instituto Gulbenkian de Ciência	2009	56	1689	DSpace	27 (17 Dez)
B-Digital	Universidade Fernando Pessoa	2006	106	1445	DSpace	831 (18 Dez)
Biblioteca Digital da FLUP	Faculdade de Letras da Universidade do Porto	2009	128	1482	Outro	279 (17 Dez)
Biblioteca Digital do IPB	Instituto Politécnico de Bragança	2007	2646	- ²⁴⁷	DSpace	682 (27 Nov)
BN Digital	Biblioteca Nacional	2006	160	-	Outro	10318 (8 Dez)
DiTeD - Depósito de Dissertações e Teses Digitais	Biblioteca Nacional	2006	292	420	Outro	2587 (20 Maio)
e-Learning Repository	Universidade do Minho	2007	459	1166	DSpace	592 (18 Dez)
Estudo Geral	Universidade de Coimbra	2008	539 1351	1271	DSpace	6696 (31 Dez)

²⁴⁶ Uma parte dos dados foi extraída do ROAR e do OpenDOAR.

²⁴⁷ O hífen significa ausência de dados.

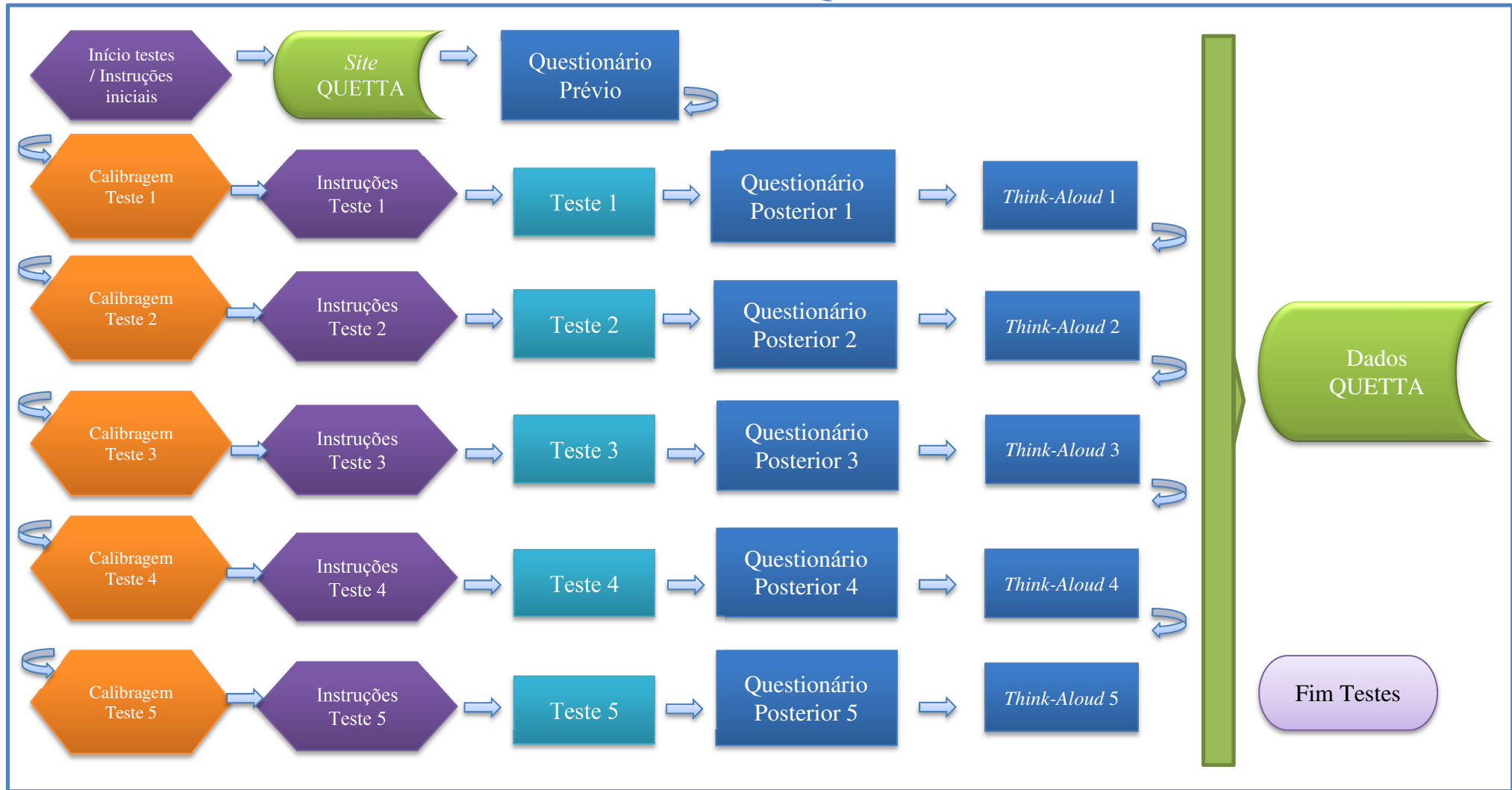
FEUP Repositório	Biblioteca da Faculdade de Engenharia da Universidade do Porto	2008	1413	1322	DigiTool	- (998 em 2 Julho de 2008)
ic-Online	Instituto Politécnico de Leiria	2009	653	1564	DSpace	148 (25 Nov)
Papadocs	Universidade do Minho	2004	1003	-	DSpace	225 (17 Dez)
PAM – Portuguese Archive of Mathematics	Portuguese Committee of the International Mathematical Union	2009	1022	-	DSpace (ID 2133)	1893 (5 Abr)
RCAAP – Repositório Científico de Acesso Aberto de Portugal	MCTES/UMIC, FCCN, UMinho	-	-	-	DSpace	34466 (18 Dez) (Agregava 25 repositórios)
REPAP – Repositório da Administração Pública	Instituto Nacional de Administração	2011	4309	2652	DSpace	3 (20 Maio)
Repositório Aberto	Universidade Aberta	2008 2012	1085 4789	1501	DSpace	660 (28 Dez)
Repositório Aberto UPorto	Universidade do Porto	2007	1086	1406	DSpace	10946 (18 Dez)
Repositório Científico da Escola Superior de Enfermagem de Coimbra	Escola Superior de Enfermagem de Coimbra	2007	6975	-	DSpace	100
Repositório Científico da Universidade de Évora	Universidade de Évora	2007 2012 2012	1088 4945 5529	1619	DSpace	95
Repositório Comum / RCAAP	RCAAP (MCTES / UMIC)	2009	2464	1714	DSpace	458 (17 Dez)
Repositório da Universidade de Lisboa	Universidade de Lisboa	2009 2010	1069 3260	1620 1961	DigiTool DSpace	1337 (12 Out) -

Repositório da Universidade dos Açores	Universidade dos Açores	2008	1089	1502	DSpace	244 (17 Dez)
Repositório Institucional dos Hospitais da Universidade de Coimbra	Centro Hospitalar e Universitário de Coimbra	2008	1098	1504	DSpace	571 (17 Dez)
Repositório da Universidade Nova de Lisboa	Universidade Nova de Lisboa	2006	1095 3353	1437	DSpace	786 (18 Dez)
Repositório da UTAD	Universidade de Trás-os-Montes e Alto Douro	2009	1090	1628	DSpace	244 (21 Set)
Repositório Digital da Universidade da Madeira	Universidade da Madeira	2009	1091	1621	DigiTool	10 (24 Jul)
Repositório do HFF	Hospital Prof. Doutor Fernando Fonseca, EPE	2010 2012	2462 6102	1715	DSpace	15 (18 Dez)
Repositório do Instituto Politécnico de Castelo Branco	Instituto Politécnico de Castelo Branco	2009	2466	1716	DSpace	100 (2 Dez)
Repositório do ISCTE	Instituto Superior de Ciências do Trabalho e da Empresa / Instituto Universitário de Lisboa	2006	1093	1199	DSpace	-
Repositório do ISPA	Instituto Superior de Psicologia Aplicada	2010	3468	2018	DSpace	41 (31 Dez)
Repositório do LNEG	Laboratório Nacional de Energia e Geologia	2009	3469	2019	DSpace	-
Repositório Institucional da ESEPF	Escola Superior de Educação Paula Frassinetti	2008	1094	1630	DSpace	271 (23 Dez)

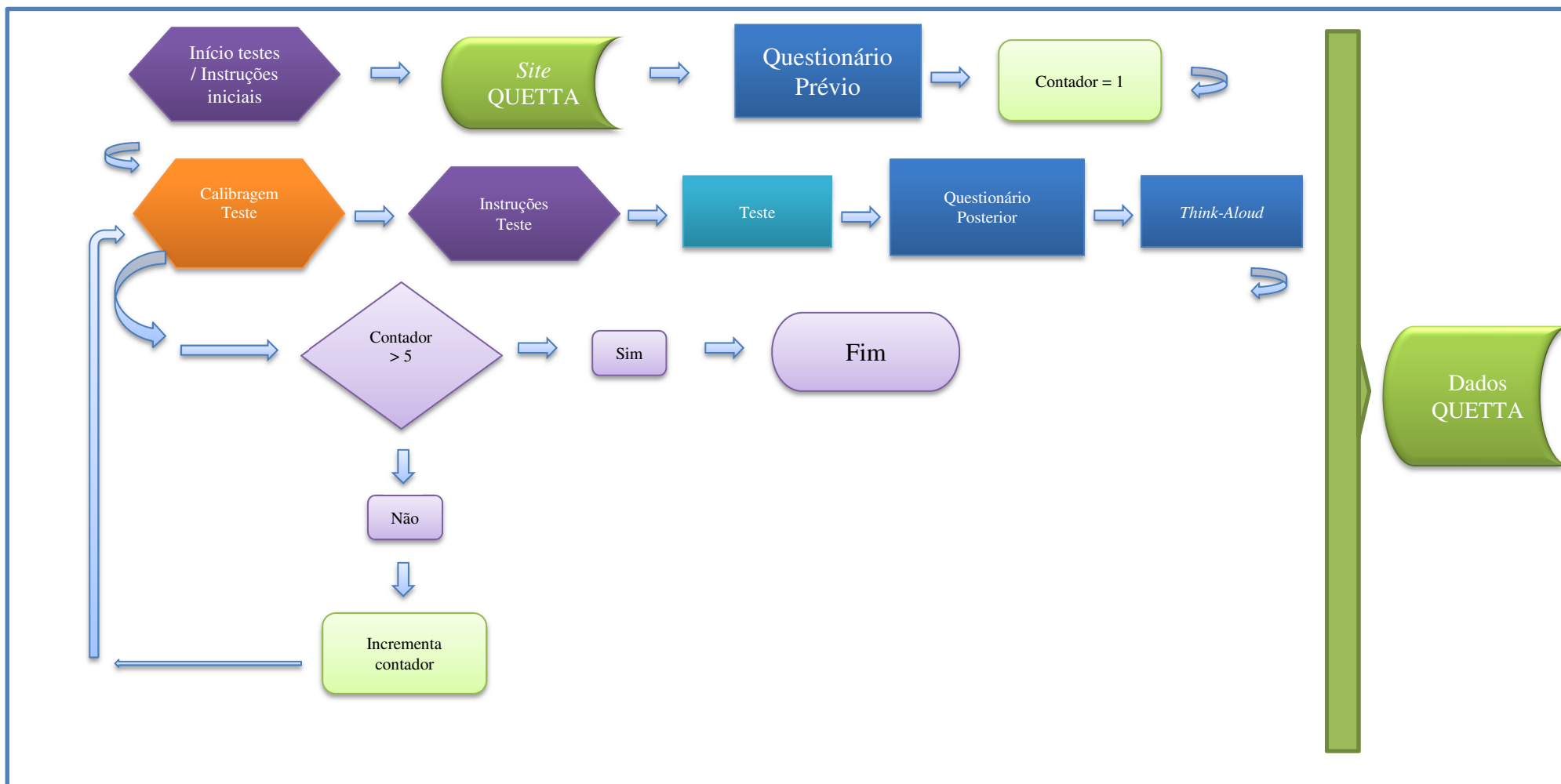
SAPIENTIA – Repositório Institucional Universidade do Algarve	Universidade do Algarve	2009	1153	1505	DSpace	170 (17 dez)
SciELO Portugal	MCTES	2006	1182	616	SciELO	17 títulos de publicações periódicas (2 de Fev)
Ubi Thesis – Conhecimento Online	Universidade da Beira Interior	2009	1319	1765	DSpace	1 (22 Dez - teste)
RepositoriUM – Universidade do Minho	Universidade do Minho	2004	1353	307	DSpace	8756 (17 Dez)
UTL Repository	Universidade Técnica de Lisboa	2008	1468	1503	DSpace	1569 (18 Dez)
WildRepositoriu m	Sociedade Portuguesa de Vida Selvagem	2007	1527	-	EPrints	-

Anexo 4

Modelo do Sistema QUETTA



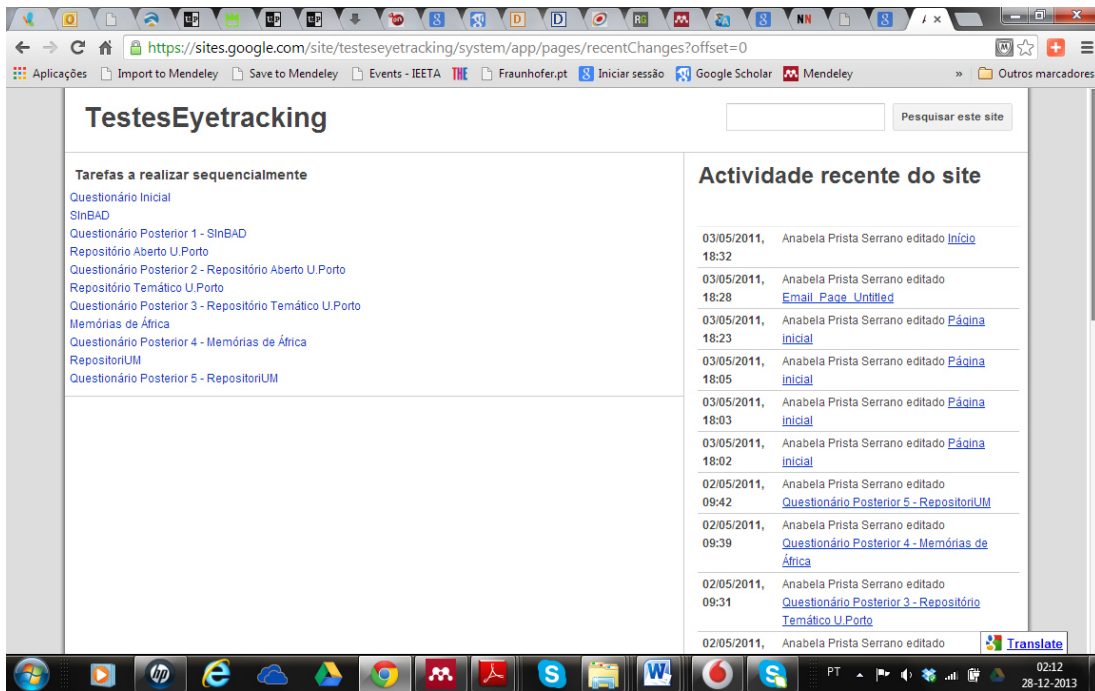
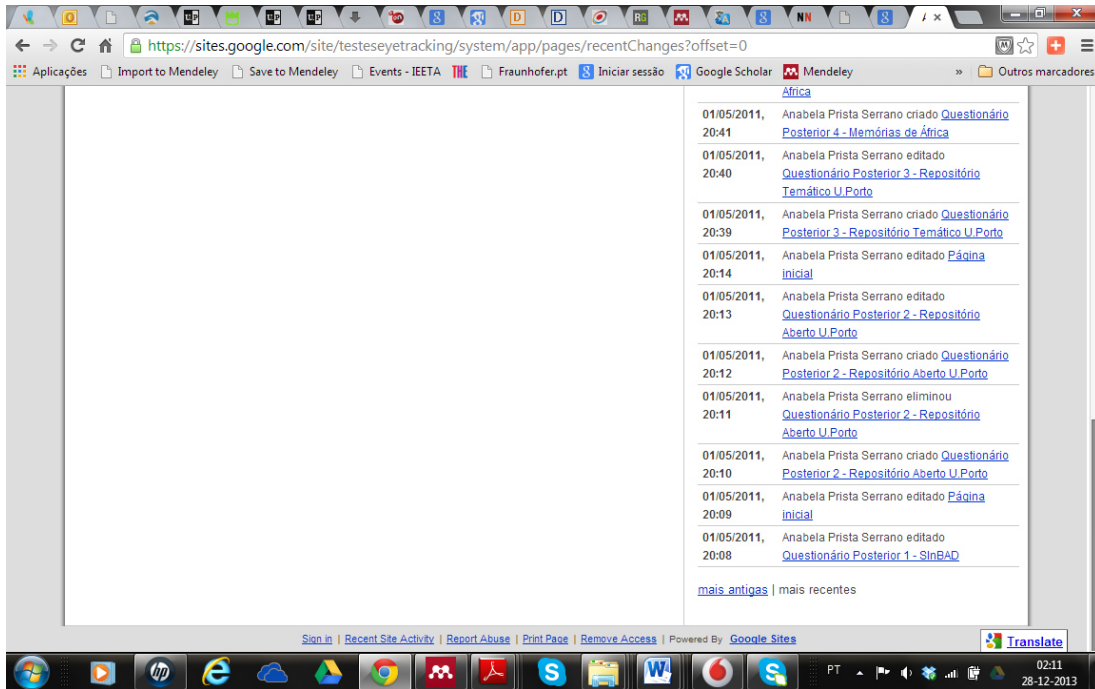
Modelo do Sistema QUETTA iniciando um contador



Anexo 5

Atividade do site QUETTA

The screenshot shows a web browser window displaying a Google Sites page. The page title is "TestesEyetracking". The main content is divided into two columns. The left column, titled "Tarefas a realizar sequencialmente", lists several tasks: "Questionário Inicial", "SInBAD", "Questionário Posterior 1 - SInBAD", "Repositório Aberto U.Porto", "Questionário Posterior 2 - Repositório Aberto U.Porto", "Repositório Temático U.Porto", "Questionário Posterior 3 - Repositório Temático U.Porto", "Memórias de África", "Questionário Posterior 4 - Memórias de África", "RepositoriUM", and "Questionário Posterior 5 - RepositoriUM". The right column, titled "Actividade recente do site", shows a log of recent changes: "01/05/2011, 20:07 Anabela Prista Serrano criou Questionário Posterior 1", "01/05/2011, 20:07 Anabela Prista Serrano editado Página inicial", "01/05/2011, 20:04 Anabela Prista Serrano editado Questionário Prévio", "01/05/2011, 19:57 Anabela Prista Serrano criou Questionário Prévio", and "01/05/2011, 19:47 Anabela Prista Serrano editado Página inicial". At the bottom of the page, there are links for "Sign in", "Recent Site Activity", "Report Abuse", "Print Page", and "Remove Access", along with a "Powered By Google Sites" notice. The browser's address bar shows the URL "https://sites.google.com/site/testeseyetracking/system/app/pages/recentChanges?offset=25". The Windows taskbar at the bottom indicates the time is 02:10 on 28-12-2013.



Anexo 6

Questionários

Questionário prévio ao Teste de Usabilidade

ID: User__

7. Qual é a sua atividade na Universidade de Aveiro? (Pode escolher mais do que uma opção)

Docente/Investigador
Profissional de Informação
Aluno Doutoramento
Aluno Mestrado/ Pós-Graduação
Aluno Licenciatura/Mestrado Integrado

8. Qual é o seu grupo etário?

18-29 anos
30-39 anos
40-49 anos
50-59 anos
60-69 anos
70-79 anos
+ de 80 anos

9. Qual é o seu nível de habilitações académicas?

Ensino secundário (12º ano)
Ensino Pós-Secundário (CET)
Bacharelato
Licenciatura
Mestrado
Doutoramento

Outro. Qual? _____

10. Costuma fazer pesquisa de informação por assunto em plataformas digitais?
(Se respondeu “não” o questionário prévio termina aqui)

Sim
Não

11. Qual a frequência com que acede a esses recursos de informação?

Frequentemente
Ocasionalmente

12. A pesquisa por assunto é para si mais usual do que a pesquisa por autor e título?

Sim
Não

Questionário posterior ao Teste de Usabilidade 1

ID: User__

Recurso de Informação - SInBAD:

<http://biblioteca.sinbad.ua.pt/Pesquisa/?sub=4>

6. Considera que foi fácil aceder à informação pretendida?

Sim
Não

7. Encontrou auxílio para a pesquisa por assunto na interface de pesquisa?

Sim
Não

8. Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa? (Pode escolher mais do que uma opção)

Índice de assuntos
Classificação
Nuvem de assuntos
Ontologia
Taxonomia

Tesouro
Nenhum

9. Qual o seu grau de satisfação com a utilização deste recurso de informação?

Muito satisfeito(a)
Satisfeito(a)
Algo satisfeito(a)
Indiferente
Algo insatisfeito(a)
Insatisfeito (a)
Muito insatisfeito(a)

10. Considera que a presença de índices por assunto, ontologias, tesouros, classificações, taxonomias, etc., facilita a realização da pesquisa por assunto?

Sim
Não
Outro _____

Questionário posterior ao Teste de Usabilidade 2

ID: User__

Recurso de Informação – Repositório Aberto U.Porto: <http://repositorio-aberto.up.pt/>

1. Considera que foi fácil aceder à informação pretendida?

Sim
Não

2. Encontrou auxílio para a pesquisa por assunto na interface de pesquisa?

Sim
Não

3. Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa? (Pode escolher mais do que uma opção)

Índice de assuntos

Classificação
Nuvem de assuntos
Ontologia
Taxonomia
Tesauro
Nenhum

4. Qual o seu grau de satisfação com a utilização deste recurso de informação?

Muito satisfeito(a)
Satisfeito(a)
Algo satisfeito(a)
Indiferente
Algo insatisfeito(a)
Insatisfeito (a)
Muito insatisfeito(a)

5. Considera que a presença de índices por assunto, ontologias, tesauros, classificações, taxonomias, etc., facilita a realização da pesquisa por assunto?

Sim
Não
Outro _____

Questionário posterior ao Teste de Usabilidade 3

ID: User__

Recurso de Informação – Repositório Temático U.Porto: <http://repositorio-tematico.up.pt/>

1. Considera que foi fácil aceder à informação pretendida?

Sim
Não

2. Encontrou auxílio para a pesquisa por assunto na interface de pesquisa?

Sim
Não

3. Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa? (Pode escolher mais do que uma opção)

Índice de assuntos
Classificação
Nuvem de assuntos
Ontologia
Taxonomia
Tesauro
Nenhum

4. Qual o seu grau de satisfação com a utilização deste recurso de informação?

Muito satisfeito(a)
Satisfeito(a)
Algo satisfeito(a)
Indiferente
Algo insatisfeito(a)
Insatisfeito (a)
Muito insatisfeito(a)

5. Considera que a presença de índices por assunto, ontologias, tesauros, classificações, taxonomias, etc., facilita a realização da pesquisa por assunto?

Sim
Não
Outro _____

Questionário posterior ao Teste de Usabilidade 4

ID: User__

Recurso de Informação – Memórias de África: <http://memoria-africa.ua.pt/introduction/tabid/83/language/pt-PT/Default.aspx>

1. Considera que foi fácil aceder à informação pretendida?

Sim
Não

2. Encontrou auxílio para a pesquisa por assunto na interface de pesquisa?

Sim

Não

3. Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa? (Pode escolher mais do que uma opção)

Índice de assuntos

Classificação

Nuvem de assuntos

Ontologia

Taxonomia

Tesouro

Nenhum

4. Qual o seu grau de satisfação com a utilização deste recurso de informação?

Muito satisfeito(a)

Satisfeito(a)

Algo satisfeito(a)

Indiferente

Algo insatisfeito(a)

Insatisfeito (a)

Muito insatisfeito(a)

5. Considera que a presença de índices por assunto, ontologias, tesouros, classificações, taxonomias, etc., facilita a realização da pesquisa por assunto?

Sim

Não

Outro _____

Questionário posterior ao Teste de Usabilidade 5

ID: User__

Recurso de Informação – RepositoriUM: <http://repositorium.sdum.uminho.pt/>

1. Considera que foi fácil aceder à informação pretendida?

Sim

Não

2. Encontrou auxílio para a pesquisa por assunto na interface de pesquisa?

Sim

Não

3. Qual destes instrumentos de apoio à pesquisa por assunto conseguiu identificar na interface de pesquisa? (Pode escolher mais do que uma opção)

Índice de assuntos

Classificação

Nuvem de assuntos

Ontologia

Taxonomia

Tesouro

Nenhum

4. Qual o seu grau de satisfação com a utilização deste recurso de informação?

Muito satisfeito(a)

Satisfeito(a)

Algo satisfeito(a)

Indiferente

Algo insatisfeito(a)

Insatisfeito (a)

Muito insatisfeito(a)

5. Considera que a presença de índices por assunto, ontologias, tesouros, classificações, taxonomias, etc., facilita a realização da pesquisa por assunto?

Sim

Não

Outro _____