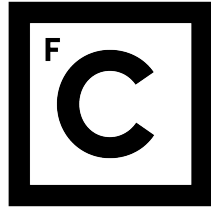


UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMÁTICA



Ciências
ULisboa

INTERFACES AVANÇADAS PARA SOLUÇÕES DE GESTÃO ARQUIVÍSTICA

Afonso Manuel Gonçalves Esteves

Mestrado em Engenharia Informática

Trabalho de projeto orientado por:
Prof.^a Doutora Diana Filipa de Pinho Costa
Prof. Doutor Wellington De Oliveira Júnior

2023

Agradecimentos

A concretização deste projeto não teria sido possível sem o apoio e a contribuição de várias pessoas e instituições, às quais gostaria de expressar a minha profunda gratidão.

Em primeiro lugar, gostaria de agradecer à Direção-Geral do Livro, dos Arquivos e das Bibliotecas, por abrir as portas a este projeto e por reconhecer a importância de melhorar o acesso aos valiosos documentos que custodiam. A sua colaboração e orientação foram fundamentais ao longo de todo o processo.

Quero estender os meus agradecimentos à *Caixa Mágica Software* por me ter inserido num projeto de tanto relevo e ainda a todos os membros da equipa de desenvolvimento que dedicaram inúmeras horas de trabalho árduo e dedicação a este projeto. Cada um desempenhou um papel crucial na conceção e desenvolvimento da interface de fácil utilização que é o cerne deste empreendimento.

Não posso deixar de expressar a minha gratidão aos utilizadores que participaram nos testes e forneceram feedback valioso. A sua contribuição foi essencial para refinar a interface e torná-la verdadeiramente centrada no utilizador.

Além disso, quero agradecer à minha família pelo apoio incondicional ao longo deste percurso académico. A vossa compreensão e encorajamento foram pilares fundamentais para alcançar este objetivo.

Por último, mas não menos importante, gostaria de agradecer aos meus orientadores que me guiaram e partilharam os seus conhecimentos ao longo desta jornada académica.

Este projeto representa um esforço coletivo e colaborativo, e é com profundo apreço que reconheço a contribuição de todos os envolvidos. Espero sinceramente que esta interface de pesquisa de arquivos revolucione a forma como interagimos com o nosso rico património arquivístico, democratizando o acesso ao conhecimento e promovendo a cultura e a pesquisa em Portugal.

Muito obrigado a todos.

Resumo

A gestão de arquivos é uma necessidade global, e Portugal não é exceção a esta necessidade. A Direção dos Arquivos Nacionais em Portugal tem a responsabilidade de custodiar uma coleção impressionante de mais de 60 milhões de documentos arquivados. Como guardião dessa rica herança histórica e cultural, é dever da Direção garantir que esses documentos estejam prontamente acessíveis tanto para a população em geral quanto para os investigadores em formato digital.

Atualmente, o acesso a esses recursos arquivísticos é facilitado através de um website; no entanto, esta porta digital não proporciona uma experiência de utilizador (UX/UI) otimizada. É evidente que a interface existente carece do design intuitivo necessário para acomodar utilizadores de diferentes origens e níveis de conhecimento, dificultando assim a sua capacidade de recuperar eficazmente os documentos que procuram. Além disso, a funcionalidade de pesquisa na plataforma atual depende de técnicas de pesquisa rudimentares, tornando o processo de localizar documentos específicos uma tarefa demorada.

O objetivo central desta tese é abordar de forma abrangente esses desafios. Este projeto visa desenvolver uma interface de fácil utilização que revolucionará a forma como os utilizadores interagem com materiais arquivísticos. Este empreendimento engloba todo o espectro do processo de design, incluindo design centrado no utilizador, metodologias de teste e a utilização de tecnologias de ponta.

Após a conclusão bem-sucedida deste projeto, é esperada a entrega de uma solução sólida de *front-end* para a plataforma web de pesquisa de arquivos em Portugal. Esta conquista marcará um marco crucial na melhoria da acessibilidade ao rico património arquivístico de Portugal, democratizando o conhecimento e promovendo a pesquisa e a exploração cultural. A plataforma idealizada capacitará utilizadores de diversas origens, incluindo académicos, historiadores, estudantes e o público em geral, para navegar e explorar sem esforço os tesouros arquivais de Portugal, preservando e promovendo o legado cultural da nação.

Palavras-chave: Arquivos, Acessibilidade Digital, Interface de Utilizador Amigável, Otimização de Pesquisa, Património Cultural

Abstract

Archival management is a global imperative, and Portugal is no exception to this pressing need. The Directorate of National Archives in Portugal is entrusted with the custodianship of an astonishing collection of over 60 million stored documents. As the custodians of this rich historical and cultural heritage, it is incumbent upon the Directorate to ensure that these documents are readily accessible to both the general population and researchers in a digital format.

Presently, access to these invaluable archival resources is facilitated through a website; however, this digital gateway falls short in providing an optimal user experience (UX/UI). It is clear that the existing interface lacks the intuitive design necessary to accommodate users of varying backgrounds and expertise levels, thereby hindering their ability to efficiently retrieve the documents they seek. Moreover, the search functionality on the current platform relies on rudimentary querying techniques, making the process of locating specific documents a daunting and time-consuming task.

The core objective of this thesis is to address these challenges comprehensively. Our research aims to develop a user-friendly interface that will revolutionize the way users interact with archival materials. This endeavor encompasses the entire spectrum of the design process, including user-centered design, rigorous testing methodologies, and the utilization of cutting-edge technologies.

Upon successful completion of this project, we anticipate delivering a robust front-end solution for the archive research web platform in Portugal. This achievement will mark a pivotal milestone in enhancing accessibility to Portugal's rich archival heritage, democratizing knowledge, and fostering research and cultural exploration. The envisioned platform will empower users from diverse backgrounds, including scholars, historians, students, and the general public, to seamlessly navigate and explore Portugal's archival treasures, thereby preserving and promoting the nation's cultural legacy.

Keywords: Archival Management, Digital Accessibility, User-Friendly Interface, Search Optimization, Cultural Heritage

Conteúdo

Lista de Figuras	xiii
Lista de Tabelas	xv
1 Introdução	1
1.1 Resumo do trabalho desenvolvido	2
1.2 Contribuições	3
1.3 Estrutura do documento	4
2 Objetivos	7
2.1 Objetivos gerais	7
2.2 Objetivos específicos	7
2.2.1 Desenvolvimento do <i>front-end</i>	7
2.2.2 Colaboração com a Equipa de UI/UX	7
3 Planeamento	11
4 Background	13
4.1 Tecnologias	13
4.1.1 React	13
4.1.2 Next.js	15
4.1.3 Nginx	16
4.1.4 Figma	16
4.1.5 Maze	17
4.1.6 Gitlab	17
4.1.7 Postman	18
4.1.8 Trello	19
4.1.9 Docker	20
4.1.10 Kubernetes	20
4.2 Conceitos	21
4.2.1 CIDOC-CRM	21
4.2.2 RDFa	21

4.2.3	ISAD(G)	22
4.2.4	UI/UX	22
4.3	Desenvolvimento ágil	24
4.4	OAI-PMH	24
5	Metodologia	27
6	Trabalho relacionado	31
6.1	Museu da Pessoa	31
6.2	Europeana	32
7	Arquitetura	35
8	Trabalho realizado	39
8.1	UI/UX aplicado ao projeto	39
8.1.1	Testes iniciais	39
8.1.2	Testes Conservação e Restauro	41
8.2	Frontoffice	45
8.2.1	Inicialização	45
8.2.2	Página inicial e login	49
8.2.3	Pesquisa Simples	49
8.2.4	Detalhes do Documento	50
8.2.5	Visualizador de Imagens	50
8.2.6	Detalhes das entidades	51
8.2.7	Árvore hierárquica do Documento	51
8.2.8	Estrutura de pastas	52
8.3	Backoffice	53
8.3.1	Dashboards	54
8.3.2	Permissões	54
8.3.3	Ambiente de Trabalho	55
8.3.4	Descrição do Documento	56
8.3.5	Associações	56
8.3.6	Intervenções do Documento	57
8.3.7	Representações Digitais	58
8.3.8	Entidades - Ator, Evento, Lugar	58
8.3.9	Árvores	60
9	Conclusão	63
	Bibliografia	67
	Bibliografia	69

Lista de Figuras

4.1	Estrutura atómica em React [17]	14
5.1	Análise do <i>backoffice</i> da plataforma Digitalq	30
6.1	Processo de agregação de informação utilizado pela Europeia	32
7.1	Arquitetura do Projeto	36
8.1	Ecrã dos resultados de uma pesquisa, na primeira fase de iterações dos testes de usabilidade	41
8.2	Ecrã dos resultados de uma pesquisa	42
8.3	Ecrã dos resultados de uma pesquisa, com o modal dos filtros de pesquisa aberto	42
8.4	Anos de experiência dos participantes no estudo	44
8.5	Correlação entre a idade, a experiência e o tempo total de conclusão das tarefas	44
8.6	Caminho dos utilizadores para concluir a tarefa: criar uma ação de recuperação num dado documento	46
8.7	Cliques no ecrã da primeira tarefa	47
8.8	Mapa de calor da tarefa de ver detalhes de uma ação	48
8.9	Página inicial, implementada no Digitalq	49
8.10	Página dos detalhes de um documento	51
8.11	Página de visualização de imagens de um documento, no <i>frontoffice</i>	52
8.12	Página dos documentos guardados por pastas	53
8.13	Ecrã da área de trabalho subdividida pelas diferentes secções	55

Lista de Tabelas

3.1	Planeamento inicial	12
3.2	Estimativa real de tempo gasto por tarefa	12
4.1	Opções de bibliotecas React open-source para a visualização de imagens	15

Capítulo 1

Introdução

A Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB), é um serviço central da administração direta do Estado Português, que tem como objetivo, assegurar a coordenação do sistema nacional de arquivos. Ao longo dos últimos anos, a DGLAB desenvolveu diversos sistemas de informação destinados a gerir informação do universo dos arquivos. Destes, o *Digitarq* assume-se como central na área de atuação da DGLAB por se destinar a gerir dados relativos ao acervo dos arquivos da rede. Tem como principais funcionalidades:

- Descrição arquivística;
- Caracterização de ações de conservação e restauro da documentação;
- Caracterização de projetos de transferência de suporte e gestão das representações digitais que deles resultam;
- Gestão de espaços físicos de depósitos.

A atual plataforma foi sendo desenvolvida ao longo do tempo, seguindo normas internacionais e orientações da própria DGLAB que foram sofrendo alterações durante o processo de uso e que condicionaram opções de recolha e processamento da informação, deixando o sistema dependente da utilização em áreas específicas de atuação da DGLAB.

Neste tipo de plataformas, com um elevado número de dados e com uma estrutura que se vai adaptando às necessidades, seguindo um paradigma de base de dados relacional, é natural surgirem problemas de incongruência e escalabilidade. A incongruência ocorre quando as mudanças nas normas e diretrizes afetam a estrutura existente da base de dados, levando a inconsistências nos dados armazenados. Por outro lado, a escalabilidade torna-se um desafio à medida que a quantidade de dados continua a crescer, demandando uma infraestrutura robusta capaz de suportar eficientemente a expansão do sistema.

Surgiu então a necessidade de existir uma versão atualizada da plataforma *Digitarq*, que modernize a forma como se interage com os arquivos a nível nacional. Para além disto, que preencha todos os requisitos da direção geral, incluindo uma nova forma de representar os dados, substituindo uma convencional base de dados relacional, por uma de grafos; um modelo de aprendizagem automática que detete atores, eventos, lugares, no modelo CIDOC-CRM [11]; uma nova

interface que acompanhe os tempos modernos e esteja adequada a todo o tipo de utilizadores. Para além de melhorar a experiência tanto para arquivistas como para investigadores, um dos principais objetivos deste projeto é torná-lo acessível e atrativo para o público em geral que, na sua maioria, não recorre à plataforma online para obter mais informações sobre arquivos, por desconhecimento ou dificuldade de usar a plataforma atual.

Isto implica num esforço considerável para garantir que a interface gráfica seja o mais intuitiva possível, com o intuito de tornar a plataforma mais amigável e de fácil navegação para aqueles que não estão familiarizados com sistemas de arquivo. O propósito é ultrapassar as barreiras de entrada e transformar a exploração dos arquivos online numa experiência gratificante e descomplicada, incentivando um público mais alargado a interagir com o património arquivístico.

Para além destes requisitos, espera-se ainda uma adaptação desta interface para o uso da web semântica, fazendo uso de dados estruturados para anotar entidades e propriedades no HTML, o que irá ajudar os motores de busca a encontrar melhores resultados fora da plataforma e permitir a sua integração numa rede mundial de conhecimento e representação arquivística.

1.1 Resumo do trabalho desenvolvido

O trabalho realizado ao longo deste projeto pode ser dividido em dois ramos principais, o *backoffice* - termo usado para referir a parte do sistema utilizado por pessoas internas à DGLAB para a gestão de informação, e o *frontoffice* - termo utilizado para referir a parte da plataforma visível aos utilizadores ou ao público em geral. Na primeira etapa do desenvolvimento, foi realizado um foco maior no *frontoffice*, em que foram realizados testes com utilizadores e foram desenvolvidos 6 dos 11 ecrãs que compõem a experiência de utilização por parte de utilizadores externos. Nesta fase deu-se o início do projeto e como tal, do ponto de vista aplicacional de *front-end*, tiveram de ser desenvolvidos inúmeros componentes de raiz, seguindo a estrutura atómica (referenciada na secção 4.1), que fossem mais tardes utilizados ao longo do desenvolvimento. Aqui foi importante uma boa comunicação com a equipa de UI/UX para garantir que o design estava a ser respeitado desde início e que estavam a ser usadas boas práticas para um sistema escalável, do ponto de vista do *design system* - conjunto consistente de diretrizes, componentes e padrões de design que garantem uma aparência e experiência coerentes em todas as páginas e interações do site. Nesta fase foi também essencial a integração do *front-end* com a equipa de dados e *back-end* para refletir a estrutura de dados complexa, resultante do uso do modelo CIDOC-CRM e do uso de três paradigmas diferentes de base de dados.

A segunda fase foi aquela que trouxe as tarefas mais complexas devido às funcionalidades exigidas. Transformar uma aplicação desktop para uma experiência web foi tão desafiante do ponto de vista do design, como da implementação. Foi necessário desenvolver um sistema de tabs e menus de navegação que guardassem também com um sistema de *auto-save* - recurso que automaticamente guarda as alterações feitas num documento - todo o conteúdo apresentado no ecrã do utilizador, para permitir uma experiência fluída ao longo de sessões e que recriasse um ambiente de aplicação desktop. Para além deste sistema, foi implementado um conjunto de funcionalida-

des para os variados grupos de utilizadores internos que possuíssem ainda uma validação de permissões. Os diversos formulários que compõem o *backoffice* e que são utilizados para descrever documentos, entidades e outros tipos de objetos, foram a maior dificuldade enfrentada. Embora utilizássemos uma biblioteca de React que facilita o processo de implementar um formulário (referencia), tivemos bastantes dificuldades quando estes se começaram a tornar mais complexos. A complexidade veio da necessidade de ter subníveis entre os campos e a possibilidade de adicionar e remover campos dinamicamente. De maneira a manter todos os formulários em concordância, montámos um template que tratava da geração dos campos e dos seus diversos estilos e particularidades a partir de uma estrutura rígida definida num objeto apenas.

1.2 Contribuições

Nesta secção, destacam-se as contribuições das várias equipas do projeto, realçando o papel crucial de cada uma no seu êxito, e ainda, as contribuições que este projeto teve a nível científico.

A equipa de UI/UX foi responsável por conceber e desenhar a interface da nova plataforma, visando uma experiência de uso amigável. Realizou ainda testes com utilizadores para ajustes iterativos, melhorando a interação. A equipa de dados implementou as bases de dados Neo4j e Elasticsearch, com uma API eficaz para comunicação com o *front-end*. Além disso, criou algoritmos de machine learning para extrair dados das base de dados já existentes. A equipa de *back-end* criou a infraestrutura e uma API para ligar o *front-end* à base de dados relacional. Geriu também a migração das bases de dados.

O Project Manager desempenhou um papel crucial na gestão da interação com o cliente, compreendendo as expectativas e necessidades do cliente, e garantindo a compreensão e implementação adequada dos requisitos. O Tech Lead supervisionou a equipa, garantindo qualidade técnica e auxiliando em questões complexas. A equipa de aplicação desenvolveu o *front-end* e participou nos testes com utilizadores, melhorando a experiência.

O projeto Digitarq representa um marco significativo no campo arquivístico, introduzindo uma série de contribuições valiosas que refletem a modernidade e a inovação tecnológica. Ao reformular a plataforma antiga com base em princípios contemporâneos e tecnologias emergentes, o Digitarq contribuiu de maneira notável para o avanço da gestão de documentos e informações num contexto global.

Um aspeto particularmente notável é a utilização de tecnologias recentes como React para o *front-end*, Neo4j para bases de dados de grafos e Elasticsearch para pesquisa avançada. Ao incorporar essas tecnologias de vanguarda, o projeto Digitarq destaca-se como um pioneiro no domínio dos arquivos, abrindo novas possibilidades de gestão eficiente de informações históricas e culturais. Essas tecnologias permitem uma experiência de utilização mais intuitiva, pesquisa eficaz e estruturação avançada dos dados arquivísticos.

A colaboração com a DGLAB realça a aplicabilidade prática das contribuições do projeto. Ao atender às necessidades específicas da instituição, o Digitarq demonstra como a modernização de plataformas arquivísticas pode ser adaptada às demandas reais das organizações culturais. Essa

colaboração também reforça a relevância das contribuições do projeto num contexto mais amplo, à medida que outras instituições possam encontrar inspiração para adotar soluções tecnológicas semelhantes.

Em resumo, o projeto Digitalq não apenas modernizou a plataforma de gestão arquivística, mas também introduziu inovações tecnológicas que têm o potencial de moldar o futuro do campo. Ao abraçar a modernidade e as tecnologias emergentes, o projeto destaca-se como um exemplo notável de como a interseção entre tecnologia e arquivística pode revolucionar a forma como as informações históricas são geridas, preservadas e disponibilizadas ao público em geral.

1.3 Estrutura do documento

Ao longo deste relatório irá ser abordado mais a fundo o trabalho efetuado anteriormente, os trabalhos semelhantes a este, as várias tecnologias envolvidas no projeto, bem como o desenvolvimento do mesmo, numa fase inicial.

O restante deste documento é composto por:

- Capítulo 2 - Onde são resumidos os objetivos gerais do trabalho.
- Capítulo 3 - Aborda o planeamento inicial bem como se este foi cumprido ou quais os obstáculos enfrentados que fizeram com que este sofresse desvios.
- Capítulo 4 - Onde é abordado o trabalho realizado anteriormente para a DGLAB, por parte da CMS (Caixa Mágica Software), as tecnologias utilizadas no projeto, bem como as suas vantagens, e ainda, alguns conceitos chave para a compreensão global deste trabalho.
- Capítulo 5 - Explica como foi organizado o desenvolvimento do projeto e também a metodologia utilizada para desenvolver esta solução de interface.
- Capítulo 6 - Menciona alguns trabalhos semelhantes a este, do ponto de vista de tecnologias utilizadas e projetos com a mesma envolvência, sendo esta, a descrição arquivística.
- Capítulo 7 - Aborda a arquitetura do projeto, do ponto de vista da sua infra-estrutura.
- Capítulo 8 - Relata todo o trabalho completado até à conclusão deste projeto de tese, bem como as dificuldades apresentadas no decorrer das tarefas e alguns pormenores técnicos da implementação da aplicação.
- Capítulo 9 - Revela as conclusões finais do trabalho.

Capítulo 2

Objetivos

2.1 Objetivos gerais

O objetivo deste projeto de tese é contribuir de forma criativa e concisa para melhorar a experiência de utilização da nova plataforma de gestão arquivística do país, o Digitalq. O contributo deste trabalho irá consistir essencialmente em integrar a equipa de *front-end* e UI/UX, construindo a aplicação desde o seu início até uma fase próxima de um estado aceitável para ser colocada em produção. Após a sua entrega, o trabalho em parceria com o cliente continuará até à sua conclusão final. É importante realçar que, até esta fase, foram alcançados os marcos essenciais para o funcionamento da plataforma, com a implementação da maioria dos ecrãs e funcionalidades, tanto no *frontoffice* como no *backoffice*.

O objetivo último é contribuir para uma plataforma moderna, intuitiva e altamente funcional, que beneficie a DGLAB e a gestão arquivística nacional.

2.2 Objetivos específicos

Para além dos objetivos gerais, é relevante especificar os objetivos específicos da área de ação deste trabalho.

2.2.1 Desenvolvimento do *front-end*

Um dos objetivos deste projeto é a implementação do *front-end* da nova plataforma Digitalq utilizando a tecnologia Nextjs. O contributo deste trabalho passa por traduzir os requisitos funcionais e estéticos em componentes interativos e reutilizáveis. O foco é garantir que a plataforma seja altamente funcional e proporcione uma experiência de utilizador envolvente.

2.2.2 Colaboração com a Equipa de UI/UX

Em paralelo com o desenvolvimento, outro objetivo específico é a colaboração ativa e produtiva com a equipa de UI/UX. Esta cooperação visa aperfeiçoar a interface e os fluxos de navegação, visando uma usabilidade otimizada. Participar nos testes de usabilidade da interface desenvolvida pela equipa de UI/UX é um elemento chave desta colaboração. Assegurar que a interface seja

intuitiva, livre de obstáculos e alinhada com as necessidades dos utilizadores é fundamental para o sucesso do projeto.

Capítulo 3

Planeamento

O plano de trabalho inicial, representado na figura 3.1, reflete muito pouco do trabalho que foi desenvolvido até à conclusão deste projeto de tese. Isto deve-se ao facto de no planeamento inicial ser apenas considerada a implementação do *front-office* da plataforma, o que não foi uma realidade. Após a escrita do relatório inicial, verificou-se a necessidade de este trabalho de tese se focar também na realização do *back-office*, uma vez que este segmento iria ser de grande esforço por parte da equipa devido às suas extensas funcionalidades, abordadas mais à frente neste trabalho.

Assim sendo, está representado numa segunda tabela, 3.2, uma estimativa mais real do que foi o desenvolvimento desta fase do projeto. De realçar os seguintes aspetos:

- A pesquisa avançada não fez parte do desenvolvimento deste trabalho, por ainda estarem a ser discutidas questões acerca do modelo de dados, que se encontra em desenvolvimento e a sofrer alterações regularmente.
- A página de ajuda não foi também desenvolvida, por se considerar uma funcionalidade com menor grau de relevância para o decorrer do projeto.
- A web semântica foi um aspeto que não foi trabalhado durante esta fase do desenvolvimento do projeto, devido a, mais uma vez, não ser considerada uma das funcionalidades mais importantes.

Embora estas funcionalidades não tenham sido desenvolvidas neste projeto de tese, irão estar presentes na versão final do projeto, por serem requisitos da aplicação.

Tarefa	Tempo estimado (dias)
Filtros e pesquisa	30
Pesquisa Avançada	16
Visualizador de Imagens	16
Página de Ajuda	8
Páginas das entidades	16
Web Semântica	20
Paginação	16
Acessibilidade	8
Melhoramentos finais	20
Escrita do relatório final	30

Tabela 3.1: Planeamento inicial

Tarefa	Tempo estimado (dias)
Filtros e pesquisa	6
Visualizador de Imagens	8
Páginas das entidades	6
Paginação	6
Backoffice - Dashboard	4
Backoffice - Permissões	6
Backoffice - Ambiente de trabalho	32
Backoffice - Descrição do Documento	35
Backoffice - Associações	16
Backoffice - Intervenções do Documento	35
Backoffice - Formulários entidades	18
Acessibilidade	8
Escrita do relatório final	30

Tabela 3.2: Estimativa real de tempo gasto por tarefa

Capítulo 4

Background

Atualmente a DGLAB possui um website chamado Digitarq, que fornece um meio dos utilizadores acederem a documentos que estejam disponíveis para consulta online ou presencial e dá ainda aos investigadores uma ferramenta essencial às suas pesquisas. Contudo, esta mesma ferramenta está em funcionamento desde 2007, e, deste modo, encontra-se desatualizada relativamente às suas tecnologias. Lado a lado com este projeto, a Caixa Mágica Software desenvolveu no ano de 2021 um website chamado CRAV, Consulta Real em Ambiente Virtual, que consiste numa plataforma que permite a interação entre a DGLAB e os cidadãos, sendo possível fazer agendamentos e pedidos de consulta de documentos, entre outras funcionalidades. Neste projeto foram adotadas as mesmas tecnologias de *software* aberto, a nível de *front-end*, como React e Node.js, que servirão como princípio para a realização desta nova plataforma.

4.1 Tecnologias

Nesta secção irão ser abordadas as principais tecnologias que irão ser utilizadas no *front-end* desta nova versão do Digitarq.

4.1.1 React

React é a principal ferramenta para o desenvolvimento do *front-end* do próximo Digitarq. É uma biblioteca de Javascript criada pela empresa Meta [14], e é usada para construir interfaces interativas e aplicações web rapidamente e eficazmente. Com esta ferramenta, criam-se componentes reutilizáveis que podem ser aplicados ao longo das páginas, tornando assim o desenvolvimento mais ágil, do que comparado com Javascript puro. Aqui a programação é feita em Javascript e JSX que se designa por Javascript XML, que, nada mais é, do que uma combinação de HTML e Javascript, permitindo a inserção de conteúdo dinâmico em tags HTML.

A principal vantagem desta tecnologia é a sua responsividade, permitindo que o *DOM* (*Document Object Model*), atualize de forma automática usando um virtual DOM e comparando as alterações do real com o virtual. Desta maneira, o React consegue atualizar a página que o utilizador está a ver apenas atualizando os componentes que forem alterados, traduzindo-se numa maior eficácia devido ao facto de não ser necessário refrescar a página por completo.

Estrutura atômica

Ao desenhar aplicações em React, devido à alta utilização de componentes, surge a necessidade de manter uma boa organização desses componentes na estrutura do projeto, de forma a que este seja escalável e compreensível para todos. É aqui que entra o modelo atômico [9] que sugere organizar os componentes em cinco níveis diferentes, conforme descrito na Figura 4.1: átomo, molécula, organismo, template e página. Os átomos são, tal como o nome indica, unidades básicas, representando componentes muito simples como botões, inputs, etc. As moléculas são conjuntos de átomos, isto é, combinações de componentes básicos. Os organismos por sua vez são conjuntos de moléculas que formam uma secção da aplicação, como por exemplo, uma barra de navegação ou um rodapé. Os templates são maioritariamente conjuntos de organismos que formam uma página. Por último, as páginas são um ecossistema que permite visualizar diferentes templates.

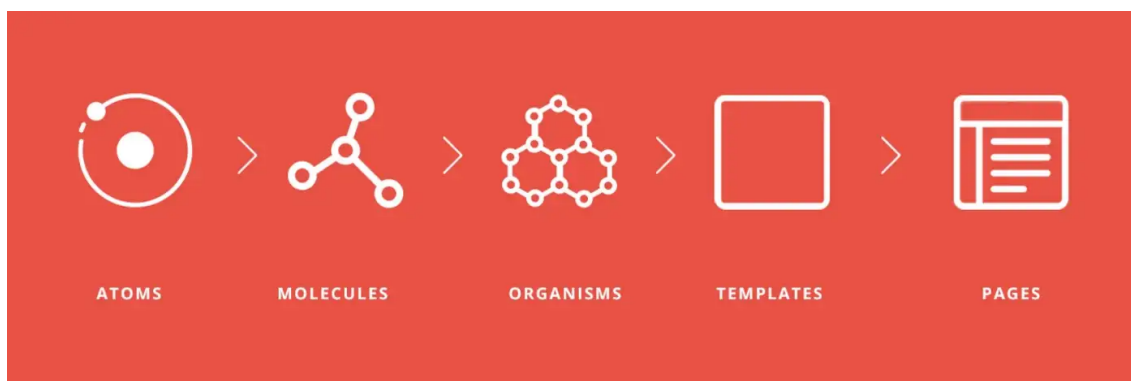


Figura 4.1: Estrutura atômica em React [17]

React Hook Form

O React Hook Form é uma biblioteca React que oferece uma maneira eficaz e flexível de lidar com formulários em aplicações web. Esta biblioteca é particularmente útil quando se trata de criar, validar e gerir formulários em projetos React. Ela baseia-se no uso de *hooks* do React para simplificar o processo de criação e manipulação de formulários. Uma das principais vantagens do React Hook Form é a sua abordagem "não controlada", o que significa que os dados do formulário são geridos pelo próprio Hook Form, eliminando a necessidade de muitos manipuladores de eventos e estados. A utilização desta ferramenta torna-se então essencial para o desenvolvimento desta plataforma que vai contar com inúmeros e complexos formulários, principalmente na parte do *backoffice*, onde será gerida toda a informação acerca dos arquivos.

React-image-pan-zoom-rotate

No âmbito de conseguir mostrar as imagens do documento, e, mais tarde, imagens no *backoffice*, foi realizada uma pesquisa de bibliotecas open-source que permitissem os formatos mais comuns

de imagens, bem como tiff. Para além deste requisito, teria de permitir fazer zoom, rodar e disponibilizar uma vista de várias miniaturas de outras imagens do documento. Rapidamente se percebeu que esta funcionalidade poderia ser realizada sem o uso de uma biblioteca, apenas com html e css. Foram várias as opções encontradas e representadas na tabela 4.1, mas optou-se pela biblioteca "react-image-pan-zoom-rotate"[10], por ser bastante simples e ter as funcionalidades necessárias.

Biblioteca	Vantagens	Desvantagens
Filerobot Image Editor	Suporta uma ampla gama de formatos de imagem, incluindo TIFF.	Pode ser complexo de configurar e usar.
TOAST UI Image Editor	Fornecer uma variedade de recursos de edição de imagem, como corte, rotação e redimensionamento.	Não é tão leve quanto outras bibliotecas.
Cropper.js	Muito fácil de usar e personalizar.	Não suporta arquivos TIFF nativamente.
React Easy Crop	Uma biblioteca de corte de imagem simples e leve.	Não suporta todos os recursos do Cropper.js.
JIMP	Uma biblioteca de processamento de imagem poderosa que pode ser usada para uma variedade de tarefas, incluindo visualização de imagens.	Pode ser complexo de usar.
React-image-pan-zoom-rotate	Uma biblioteca simples e leve para zoom, rotação e panning de imagens.	Não suporta todos os recursos de outras bibliotecas.

Tabela 4.1: Opções de bibliotecas React open-source para a visualização de imagens

4.1.2 Next.js

Next.js é uma framework de React que permite melhorar o desempenho do carregamento de páginas web, disponibilizando três diferentes tipos de renderização do conteúdo na página. React essencialmente renderiza o seu conteúdo no cliente, o que limita bastante a velocidade de processamento. Esta limitação em velocidade acarreta várias desvantagens em termos de experiência de utilização e *SEO (Search Engine Optimization)*, o que faz do Next.js um bom candidato para empresas que queiram ver o seu website dinâmico mas ao mesmo tempo com um bom desempenho.

Next.js usa Node.js - um ambiente de servidor open source que usa JavaScript, permitindo assim renderizar o conteúdo React no lado do servidor. Para além desta grande vantagem, esta ferramenta é flexível e permite, como opção, realizar a renderização de forma estática, em determinadas partes do site, fazendo com que a rapidez de carregamento do mesmo seja ainda mais maior, com a desvantagem de o conteúdo deixar de ser dinâmico e ter a necessidade de ser pré-carregado. Se nenhuma destas opções for ideal, o Next.js apresenta ainda uma terceira solução

que combina o melhor das duas anteriores: *Static incremental*, que permite gerar certa informação estaticamente mas ainda assim atualizar dinamicamente elementos que não tenham sido gerados estaticamente ou que tenham sido alterados.

4.1.3 Nginx

O Nginx é um servidor web de código aberto amplamente utilizado que também funciona como proxy reverso, balanceador de carga e servidor de cache. É conhecido pelo seu desempenho, escalabilidade e capacidade de lidar com um grande volume de pedidos simultâneos. Ele recebe pedidos dos clientes, encaminha-os para os servidores apropriados e, em seguida, envia as respostas dos servidores de volta aos clientes.

Além de ser um servidor web, o Nginx é frequentemente utilizado como um proxy reverso para distribuir o tráfego entre vários servidores, melhorando a escalabilidade e a disponibilidade das aplicações web. Também pode ser configurado como um servidor de cache para armazenar em cache conteúdos estáticos, reduzindo assim a carga nos servidores de aplicação e melhorando o desempenho.

4.1.4 Figma

Na realização do projeto Digitarq, a tecnologia Figma desempenhou um papel crucial ao fazer a ponte entre a equipa de UI/UX e a equipa de *front-end*. O Figma é uma poderosa ferramenta de design colaborativo baseada na nuvem, que permite aos designers criar, partilhar e iterar sobre projetos de interface de forma eficiente e colaborativa.

A equipa de UI/UX utilizou o Figma para desenvolver os protótipos de design da nova plataforma Digitarq. Nesta etapa, os designers criaram wireframes, layouts e mockups que refletiam as ideias e as especificações de design estabelecidas. O Figma permitiu que os membros da equipa trabalhassem em conjunto, em tempo real, no mesmo projeto, possibilitando a partilha instantânea de alterações e feedback.

Essa colaboração direta entre a equipa de UI/UX e a equipa de *front-end* foi essencial para a eficiente implementação da interface da plataforma. A utilização do Figma como ponto de partida permitiu que os developers entendessem claramente as intenções de design e as interações pretendidas. Dessa forma, o processo de transformar os designs em código tornou-se mais suave e preciso, uma vez que a equipa de *front-end* dispunha de todas as informações detalhadas para a implementação.

Além disso, o Figma também ofereceu recursos para a criação de componentes reutilizáveis, o que possibilitou uma maior consistência e padronização na interface da plataforma. Os componentes reutilizáveis simplificaram o desenvolvimento, economizando tempo e esforço, e ajudaram a garantir uma experiência coesa para os utilizadores em diferentes partes da aplicação.

Em suma, a tecnologia Figma foi uma ferramenta fundamental na jornada do projeto Digitarq, ao desempenhar um papel crucial como intermediário entre a equipa de UI/UX e a equipa de *front-end*. Ao facilitar a colaboração, o Figma contribuiu para uma implementação mais eficiente

e precisa da interface da plataforma, resultando numa experiência de utilização mais consistente e satisfatória para os utilizadores.

4.1.5 Maze

O Maze é uma plataforma de testes de usabilidade que permite criar protótipos interativos, conduzir testes com utilizadores e obter valiosos insights sobre a experiência do utilizador. Na realização do projeto, esta tecnologia desempenhou um papel fundamental ao facilitar e enriquecer a segunda fase de testes com utilizadores, promovendo a colaboração entre a equipa de UI/UX e a equipa de aplicação.

Uma das principais vantagens do Maze é a sua facilidade de uso, que permitiu à equipa de UI/UX criar protótipos interativos de alta fidelidade de forma rápida e intuitiva. Esses protótipos refletiam fielmente a aparência e a interatividade da nova plataforma Digitalq, possibilitando aos utilizadores testar a interface em um ambiente controlado antes de se avançar para a sua implementação. Esta abordagem permitiu à equipa de UI/UX realizar ajustes e refinamentos com base nos feedbacks dos utilizadores, contribuindo para uma melhor adaptação às suas necessidades e expectativas.

Outra vantagem fundamental do Maze é a sua capacidade de realizar testes com os utilizadores de forma remota e escalável. Apesar desta funcionalidade estar presente na plataforma, a equipa decidiu fazer o máximo de testes possíveis presencialmente, devido à complexidade da interface e da pouca experiência dos utilizadores com websites no geral.

A tecnologia Maze também proporcionou uma análise detalhada dos dados recolhidos durante os testes. Através de métricas e análises quantitativas, a equipa pôde identificar padrões de interação, pontos problemáticos e áreas que requeriam melhorias na usabilidade. Além disso, a plataforma permitiu a gravação das sessões de teste, o que possibilitou uma revisão mais aprofundada do comportamento dos utilizadores durante a interação com o protótipo.

Em resumo, a tecnologia Maze revelou-se uma ferramenta valiosa ao auxiliar a equipa de UI/UX e a equipa de aplicação na fase de testes com os utilizadores. As suas vantagens, como a facilidade de utilização, a capacidade de realizar testes remotos e uma análise detalhada dos dados, foram fundamentais para aprimorar a usabilidade e a experiência geral da nova plataforma Digitalq. A colaboração e interação entre as equipas foram fortalecidas, resultando num projeto mais coeso, centrado no utilizador e alinhado com as necessidades do cliente.

4.1.6 Gitlab

A plataforma Gitlab desempenhou um papel fundamental na gestão e controlo do código fonte durante o projeto Digitalq. O Gitlab é uma plataforma de desenvolvimento colaborativa baseada em Git, que permite a criação, gestão e controlo de repositórios de código fonte de forma eficiente e segura.

Uma das principais vantagens da utilização do Gitlab foi a capacidade de permitir que as diferentes equipas envolvidas no projeto mantivessem o controlo do código através de diferentes

repositórios. A a equipa de aplicação, a equipa de dados e a equipa de *back-end* puderam trabalhar em conjunto de forma coordenada, cada uma com o seu próprio repositório, mantendo a independência e o controlo necessário para as suas áreas específicas de trabalho.

Além disso, o Gitlab proporcionou uma gestão eficiente dos *branches* (ramificações) de código. Cada equipa pôde criar e gerir os seus próprios branches para desenvolver e testar as suas funcionalidades sem interferir diretamente com o código em ambiente de qualidade. Essa separação permitiu um desenvolvimento mais organizado e um controlo adequado das alterações implementadas.

Além disso, o Gitlab possibilitou a implementação do processo de continuous deployment. Após os commits serem realizados manualmente para a branch principal (main), o deployment das alterações era realizado através da simples ação de fazer o pull do repositório Git nas máquinas adequadas. Isso permitia que as alterações fossem implementadas nos ambientes apropriados, facilitando os testes e verificações adicionais antes da disponibilização para qualidade (ambiente este que simula o ambiente de produção e serve para as funcionalidades serem testadas com mais rigor incluindo o cliente).

Dessa forma, a plataforma Gitlab e a prática de continuous deployment proporcionaram uma gestão eficiente do código fonte e uma implementação controlada das alterações. Através da colaboração entre as equipas e do controlo adequado do código, foi possível manter a qualidade do software e garantir uma implementação segura e eficiente das alterações desenvolvidas ao longo do projeto Digtarq.

4.1.7 Postman

A plataforma Postman desempenhou um papel fundamental no desenvolvimento do projeto Digtarq, especialmente para a equipa de aplicação. O Postman é uma plataforma de colaboração que permite testar e documentar APIs de forma eficiente.

Na equipa de aplicação, utilizámos o Postman para testar as rotas desenvolvidas pela equipa de *back-end* e dados antes de as implementar na aplicação. Esta abordagem permitiu-nos validar a correta integração entre os diferentes componentes do sistema e garantir o funcionamento adequado das APIs.

Com o Postman, conseguimos enviar requests HTTP às rotas desenvolvidas pela equipa de *back-end* e dados, fornecendo os parâmetros e dados necessários para simular cenários de utilização real. Isto permitiu-nos verificar se as respostas da API estavam corretas, se os dados eram retornados conforme o esperado e se as funcionalidades estavam de acordo com os requisitos estabelecidos.

Além disso, o Postman facilitou o teste de diferentes métodos HTTP, como GET, POST, PUT e DELETE, permitindo-nos explorar as várias funcionalidades disponíveis nas APIs. Também foi possível configurar cabeçalhos personalizados, autenticação e outros parâmetros relevantes para cada rota, garantindo a correta interação com o *back-end*.

Outra vantagem do Postman foi a possibilidade de guardar e organizar as diferentes requisições

em coleções. Isto permitiu-nos criar conjuntos de testes específicos para cada funcionalidade, facilitando a execução e reexecução dos testes ao longo do desenvolvimento.

A utilização do Postman durante o desenvolvimento proporcionou uma abordagem mais eficiente e segura para testar as rotas desenvolvidas pela equipa de *back-end* e dados. Isto permitiu-nos identificar e corrigir erros mais cedo no processo de desenvolvimento, evitando a necessidade de implementar funcionalidades defeituosas na aplicação.

4.1.8 Trello

O Trello é uma plataforma de gestão de projetos baseada em quadros (boards), listas (lists) e cartões (cards), que permite organizar tarefas, acompanhar o progresso e facilitar a colaboração entre as equipas.

A utilização do Trello proporcionou uma abordagem ágil para a gestão das atividades do projeto. As equipas envolvidas no desenvolvimento do Digitarq puderam criar cartões específicos para cada fase do projeto, o que incluiu desde o planeamento e design até a implementação e testes. O quadro foi dividido em listas, que representavam os diferentes sprints ou outros tipos de categorias como tarefas em backlog ou dúvidas para o cliente.

Dentro de cada lista, as tarefas individuais foram representadas por cartões, contendo dados detalhados sobre as atividades específicas a serem realizadas. Cada cartão podia ser atribuído a um membro da equipa responsável pela sua execução e continha os story points e descrições claras das tarefas. Isso permitiu uma gestão mais precisa dos fluxos de trabalho, evitando ambiguidades e mantendo todos os membros da equipa atualizados sobre o status das atividades em tempo real.

A utilização de “story points” nos cartões do Trello trouxe uma vantagem significativa na gestão do projeto, permitindo manter um registro mais preciso das tarefas realizadas em cada sprint. Os “story points” são uma técnica comum em metodologias ágeis, como o Scrum, que permite estimar a complexidade e o esforço necessário para concluir uma tarefa, em vez de estimar o tempo específico para a sua conclusão.

Ao adotar os “story points” nos cartões do Trello, a equipa pôde avaliar a dificuldade relativa de cada tarefa em comparação com outras, sem estar vinculada a prazos rígidos. Essa abordagem contribuiu para uma maior flexibilidade no planeamento das iterações do projeto, uma vez que os “story points” fornecem uma métrica mais abstrata, menos suscetível a variações de tempo decorrentes de fatores externos ou imprevistos.

Essa métrica subjetiva permitiu à equipa evitar uma pressão excessiva para cumprir prazos pré-determinados, focando-se, em vez disso, em garantir a qualidade das entregas. À medida que as iterações do projeto avançavam, a equipa ganhou mais experiência e precisão nas estimativas de “story points”, aprimorando a sua capacidade de prever o esforço necessário para concluir novas tarefas.

A utilização dos “story points” em conjunto com a abordagem do Trello também facilitou o planeamento e a revisão de cada sprint. Durante as reuniões de planeamento, a equipa selecionava as tarefas prioritárias para o próximo sprint com base nos “story points” associados a cada cartão.

Dessa forma, foi possível equilibrar a carga de trabalho entre os membros da equipa e garantir que o sprint fosse viável de ser concluído no período definido.

Além disso, ao final de cada sprint, durante os sprint reviews, os “story points” atribuídos às tarefas concluídas permitiram avaliar o desempenho da equipa, comparando as estimativas iniciais com o trabalho efetivamente realizado. Essa análise retrospectiva possibilitou o aperfeiçoamento contínuo do processo de estimativa e planeamento para os próximos sprints, aprimorando a previsibilidade e a eficiência do projeto.

Em resumo, a tecnologia Trello foi uma ferramenta valiosa que apoiou a gestão ágil do projeto Digitarq. Através da organização visual, da colaboração em tempo real e da implementação de práticas ágeis, o Trello facilitou a coordenação das equipas e a gestão eficiente das atividades.

4.1.9 Docker

O Docker é uma plataforma *open-source* de virtualização de contentores que revolucionou a forma como as aplicações são desenvolvidas, distribuídas e executadas. Ao adotar a tecnologia de contentores, o Docker permite que as aplicações e todos os seus componentes, como código, bibliotecas e dependências, sejam empacotados em contentores leves e autónomos.

Estes contentores são altamente portáteis, o que significa que uma aplicação embalada num contentor Docker pode ser executada em qualquer ambiente que suporte o Docker, desde o ambiente de desenvolvimento local até aos servidores na nuvem.

Além disso, o Docker oferece isolamento de recursos, o que significa que cada contentor é independente e não interfere com outros contentores no mesmo sistema. Isso torna a implementação de várias aplicações no mesmo servidor mais segura e eficiente.

A inicialização rápida dos contentores é outra vantagem significativa. Os contentores Docker podem ser iniciados em questão de segundos, tornando possível escalar rapidamente uma aplicação em resposta a picos de tráfego ou alocar recursos com eficiência.

O Docker também simplifica a gestão de contentores, oferecendo ferramentas de linha de comando e uma interface gráfica amigável. Além disso, pode ser integrado com orquestradores, como o Kubernetes, para facilitar a gestão de contentores em grande escala.

4.1.10 Kubernetes

O Kubernetes é uma plataforma *open-source* para automação, escalonamento e gestão de aplicações em contentores. Ele facilita a implementação, a gestão e a orquestração de contentores, permitindo que as aplicações sejam executadas de forma mais eficiente e fiável.

O Kubernetes é como um “gestor de tráfego” para as suas aplicações em contentores. Ajuda a garantir que as suas aplicações estejam sempre disponíveis, escaláveis e em funcionamento, mesmo quando ocorrem falhas ou quando é necessário atualizar partes do sistema. Ao automatizar tarefas de gestão, como a distribuição de carga, o escalonamento automático e a recuperação de falhas, o Kubernetes simplifica a gestão de aplicações complexas e ajuda as equipas de desenvolvimento a concentrarem-se no desenvolvimento de software, em vez de se preocuparem com a

infraestrutura subjacente.

4.2 Conceitos

Depois de definir as tecnologias utilizadas neste projeto, é ainda importante descrever alguns conceitos essenciais na realização deste projeto, que se prendem com a ontologia do novo modelo de dados.

4.2.1 CIDOC-CRM

O CIDOC-CRM (CIDOC Conceptual Reference Model) é um modelo de referência conceptual que inclui uma ontologia desenvolvida para definir conceitos e informações relacionadas com herança cultural e documentação de museus. Este modelo foi criado pelo Comité Internacional para Documentação, uma parte da organização não governamental *ICOM (International Council of Museums)*, e foi aceite como uma referência em 2006. O principal objetivo do CIDOC-CRM é disponibilizar um modelo que permita às instituições de todo o mundo descrever os seus arquivos históricos de forma uniformizada.

Ao adotar o CIDOC-CRM, as instituições podem obter uma compreensão genérica dos dados relacionados ao património cultural. O modelo fornece uma estrutura semântica comum e extensível, permitindo a descrição de qualquer tipo de documento dentro dos parâmetros definidos. Isso significa que a aplicação do CIDOC-CRM oferece uma base sólida para a representação e gestão da informação arquivística de forma consistente e interoperável.

Além disso, o CIDOC-CRM é um modelo extensível, o que significa que a plataforma Digitaltarq pode adaptar-se e expandir-se para acomodar diferentes tipos de documentos e necessidades específicas. A flexibilidade do CIDOC-CRM permite que a plataforma evolua ao longo do tempo, à medida que novos requisitos e funcionalidades são identificados. Isso garante a sua escalabilidade e capacidade de crescimento para atender às demandas em constante mudança no campo da gestão arquivística.

Em resumo, o CIDOC-CRM é um modelo de referência conceptual com uma ontologia associada, que oferece uma estrutura semântica comum e extensível para a descrição de arquivos históricos e património cultural. A sua adoção pelas instituições permite uma gestão consistente e interoperável da informação arquivística, promovendo a partilha de dados e a colaboração entre as organizações envolvidas na preservação e estudo do património cultural.

4.2.2 RDFa

RDFa, ou, *Resource Description Framework in Attributes*, é uma recomendação da W3C, *World Wide Web Consortium*, para incorporar metadados em páginas Web. O modelo de dados RDF (*Resource Description Framework*) permite adicionar triplos sujeito, predicado, objeto em documentos escritos em XHTML, tudo numa ótica de melhorar a estruturação dos dados disponíveis na Internet. RDFa veio colmatar a necessidade existente de aplicar este modelo a páginas HTML. Con-

trariamente ao modelo JSON-LD, o RDFa permite definir propriedades e atributos nas próprias que serão adicionados aos próprios elementos HTML, ao invés de estarem definidos num elemento script, utilizado como bloco de dados. A vantagem de usar este modelo neste projeto prende-se com a necessidade de definir propriedades diretamente no HTML, provenientes do modelo de dados desenhado, que contempla a presença de entidades por meio de Strings.

4.2.3 ISAD(G)

Fundamental para compreender na plenitude este projeto, é definir o que é o *General International Standard Archival Description* (também conhecido como ISAD(G)) [13]. Este acordo define uma lista de elementos e regras para a descrição de arquivos e descreve os tipos de informação que devem ser incluídos nessas descrições. ISAD(G) segue quatro princípios essenciais como:

1. Descrições, do mais geral, para o mais específico.

As descrições com mais do que um nível seguem sempre uma estrutura hierárquica, que, normalmente começa por ser um *Fond* (Fundo), ou uma *Collection* (Coleção) - dois níveis de descrição presentes no modelo ISAD(G) e traduzidos para a língua portuguesa [4], e que, seguirá para níveis mais abaixo na hierarquia como por exemplo um subfundo, uma série ou um documento.

2. Informação relevante para o nível de descrição.

A informação que está descrita em cada nível de descrição necessita de ser específica para esse mesmo nível.

3. Relações entre descrições.

Cada unidade arquivística necessita de apresentar uma relação direta para o nível diretamente acima. Deste modo, recursivamente, consegue-se obter a árvore completa de cada unidade.

4. Não repetição de informação.

De forma a evitar repetições, as informações mais gerais necessitam estar definidas nos níveis mais acima, semelhante ao que acontece com o princípio 2.

Para além destes princípios, existem ainda vinte e seis tipos de dados que são aconselhados ter quando se descreve um arquivo, ou documento, sendo que seis deles são obrigatórios. Este conjunto de dados está dividido em sete categorias diferentes, que juntas compõem uma descrição completa da unidade arquivística. Os dados obrigatórios a estarem definidos são o código de referência, o título, a data, nível de descrição, dimensão e suporte, e o nome do criador da unidade de descrição.

4.2.4 UI/UX

Muitas questões surgem quanto à diferença existente entre estes dois conceitos, que embora distintos, se complementam um ao outro, sempre focando a atenção no utilizador.

User Interface

Interface do Utilizador engloba temas mais estéticos como teoria das cores, tipografia, padrões de design e para além disto foca-se também em criar interfaces funcionais e intuitivas. Grande parte deste trabalho é usualmente feito por designers que para além de desenhar as interfaces têm também a tarefa de prototipar, seja em versões de baixa ou alta fidelidade, que mais tarde podem ser aplicadas em testes com utilizadores.

User Experience

A experiência de utilizador caracteriza-se por ser a relação entre um utilizador e um determinado produto, ou serviço. Aqui o desafio é entender as necessidades da pessoa que vai usufruir do serviço, e entender se a interface que se está a desenhar é simples, intuitiva e que sentimentos provoca no utilizador. Aqui a parte emocional também é preciso ser tida em conta, seja um sentimento de frustração, desconforto ou qualquer outro, para serem feitos ajustes à interface de forma a conseguir transmitir uma experiência agradável para cada utilizador. É nesta categoria que encaixam os testes de usabilidade para perceber que lacunas existem no design feito até ao momento e ainda possíveis melhorias.

Métodos e estudos com utilizadores

Métodos e estudos com utilizadores é um tópico bastante relevante na área do UI/UX e na perspectiva de desenvolver uma interface ideal. Quanto aos métodos em si podemos dividi-los essencialmente em quatro tipos: observação, questionário, entrevista e teste de usabilidade.

Observação consiste em observar um utilizador a realizar determinada tarefa num cenário natural, por exemplo, pessoas a utilizar telemóveis num aeroporto. Este método apesar de ser benéfico pelo reconhecimento de padrões neste ambiente natural, exige muito tempo para que se possa obter os dados pretendidos.

Questionários são formas muito eficazes de obter dados depressa devido à facilidade com que são enviados para um número elevado de participantes, mas também podem induzir em erros caso haja perguntas menos claras.

As **entrevistas** são vantajosas pelo facto de as questões que surgem podem ser rapidamente esclarecidas por parte dos entrevistadores e pode também dar aso a outro tipo de questões que não tinham sido previamente pensadas. Por outro lado, este método exige mais tempo do que questionários e limita assim o número de participantes.

Por último, os **testes de usabilidade**, o método escolhido para ser realizado neste projeto, é um método bastante prático e económico para encontrar problemas cruciais na usabilidade desenhada. Em contra-partida, não garante que todos os problemas sejam descobertos. Em qualquer estudo é essencial definir as hipóteses de pesquisa, que nada mais é do que uma declaração precisa do problema que pode ser testado através de uma investigação empírica. Há dois tipos de hipóteses possíveis, a H0, ou hipótese nula, e a H1, ou hipótese alternativa. A hipótese nula normalmente representa que não existe diferença entre condições experimentais, já a H1, é uma declaração que

é mutuamente exclusiva com a hipótese nula. Os resultados empíricos recolhidos durante o estudo refutam ou confirmam alguma das hipóteses dentro de um determinado intervalo de confiança.

Uma hipótese deve sempre ter presentes as variáveis dependentes e independentes, sendo as dependentes aquelas que são dependentes do utilizador e que normalmente requerem a atenção de quem está a realizar o estudo. As independentes são as variáveis que não dependem do comportamento do utilizador e são condições que são controláveis pelos investigadores.

Quanto ao design do estudo, há que se ter em conta como se realizam os testes, se é um estudo *between-group* ou *within-group*. *Between* significa que, existindo vários grupos de participantes, um determinado grupo testa apenas uma condição da experiência. Isto traz vantagens como uma diminuição do tempo de participação, que conseqüentemente traz menos fadiga e frustração aos participantes. Em contra-partida, é mais difícil de observar diferenças entre as condições e requer uma amostra maior de indivíduos. Por outro lado, o *within-subjects* implica que todos os participantes testem todas as condições, o que requer uma amostra mais pequena, mas que leva a um tempo maior a terminar o estudo e que afeta a aprendizagem entre condições. Por exemplo, um participante pode ter realizado uma tarefa numa condição, aprendido com ela, e aquando da realização da mesma tarefa noutra condição, o seu desempenho pode se revelar diferente devido a já a ter realizado anteriormente. Desta maneira, o resultado do estudo pode sofrer alterações não previstas.

4.3 Desenvolvimento ágil

O desenvolvimento ágil é uma abordagem moderna e flexível para a gestão de projetos de software e desenvolvimento de produtos. Esta metodologia, amplamente adotada na indústria de tecnologia, promove a colaboração interativa, a adaptabilidade e a entrega contínua de valor aos clientes.

Ao contrário dos métodos de desenvolvimento tradicionais, o desenvolvimento ágil enfatiza a comunicação regular e próxima entre os membros da equipa de desenvolvimento e os clientes. As equipas trabalham em ciclos curtos e iterativos, conhecidos como *sprints*, onde definem metas claras para cada iteração e priorizam funcionalidades com base no feedback e nas necessidades em constante evolução do cliente.

O desenvolvimento ágil dá prioridade ao software funcional em detrimento de documentação extensiva e à capacidade de se ajustar às mudanças em vez de seguir um plano fixo. Isso possibilita que as equipas sejam ágeis na adaptação a alterações nos requisitos e às necessidades emergentes.

4.4 OAI-PMH

O protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) é um padrão de protocolo de comunicação usado para recolher metadados de repositórios digitais. A sua principal finalidade é facilitar a interoperabilidade e a partilha de informações entre diferentes sistemas de gestão de conteúdo, como bibliotecas digitais, arquivos e repositórios de dados.

A característica fundamental do OAI-PMH é que ele permite que um sistema, chamado “recolhedor”, recupere metadados de um ou mais repositórios de dados, conhecidos como “fornecedores”. Os metadados consistem em informações descritivas sobre recursos digitais, como títulos, autores, datas e outros detalhes relevantes.

O OAI-PMH funciona com um modelo simples de pedido e resposta. O recolhedor envia pedidos ao fornecedor, especificando o que deseja recolher (por exemplo, todos os metadados de um tipo específico de recurso) e o fornecedor responde com os metadados correspondentes num formato padronizado, como XML.

Uma das vantagens do OAI-PMH é a sua capacidade de recolher metadados de múltiplos fornecedores de forma eficiente e uniforme. Isso facilita a criação de agregadores de conteúdo, como catálogos de bibliotecas digitais, motores de busca académicos e outros serviços que pretendem reunir informações de diversas fontes diferentes.

Capítulo 5

Metodologia

Neste capítulo descreveremos a abordagem metodológica utilizada para atingir os objetivos do projeto Digitalq. O estudo adotou uma abordagem mista, combinando elementos de pesquisa qualitativa e quantitativa, a fim de obter uma compreensão abrangente da reformulação da plataforma e suas funcionalidades.

Inicialmente, a equipa de UI/UX assumiu a responsabilidade de estruturar a nova plataforma Digitalq. Para isso, foram realizadas pesquisas detalhadas sobre as melhores práticas de design de interfaces e experiência do utilizador. De realçar que os detalhes sobre as pesquisas realizadas estão fora do escopo desta tese. Foi ainda analisada a plataforma anterior e identificados os pontos fortes e as áreas que precisavam ser melhoradas. Este levantamento foi registado em forma de texto, juntamente com capturas de ecrã para melhor análise, como demonstrado na figura 5.1

Além disso, agendamos reuniões com a equipa da Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB) para obter um entendimento profundo das necessidades e requisitos específicos para a reformulação da plataforma. Durante essas reuniões, foram discutidas as funcionalidades essenciais, desenvolvidas ao longo do capítulo 8, as expectativas dos utilizadores e os objetivos gerais do projeto.

Com base nas pesquisas realizadas, análise da plataforma anterior e nas reuniões com a DGLAB, desenvolvemos uma nova estrutura e fluxo de navegação para a plataforma Digitalq. O foco foi a usabilidade, a acessibilidade e a eficiência das interações do utilizador com a plataforma.

Para validar a nova estrutura proposta, foram realizados testes de usabilidade com utilizadores, incluindo arquivistas, utilizadores externos (que não estão familiarizados com a plataforma antiga) e outros profissionais envolvidos na gestão de documentos e arquivos. Observamos o seu comportamento ao utilizar a plataforma e recolhemos *feedback* valioso sobre a experiência do utilizador. Esses testes permitiram-nos fazer ajustes na interface, garantindo que a plataforma atendesse às necessidades e expectativas dos utilizadores. Na secção 8.1 do capítulo 8 irão ser detalhados esses mesmos ajustes e o procedimento completo dos testes de usabilidade.

Com base nos dados recolhidos e nas melhores práticas identificadas na revisão da literatura, avançamos para a fase de desenvolvimento e implementação da nova plataforma Digitalq. Nesta etapa, a equipa de desenvolvimento trabalhou em conjunto com a equipa de UI/UX para concretizar os requisitos e as ideias obtidas previamente.

Durante o processo de desenvolvimento e implementação da plataforma Digitalq, adotámos um fluxo de trabalho colaborativo e iterativo. Para garantir uma comunicação eficiente e alinhamento contínuo com a equipa, realizamos reuniões diárias internas com todos os membros da equipa envolvidos no projeto. Estas reuniões diárias, conhecidas como *daily-meetings*, permitiram-nos partilhar atualizações sobre o progresso, identificar possíveis obstáculos e alinhar as tarefas para o dia.

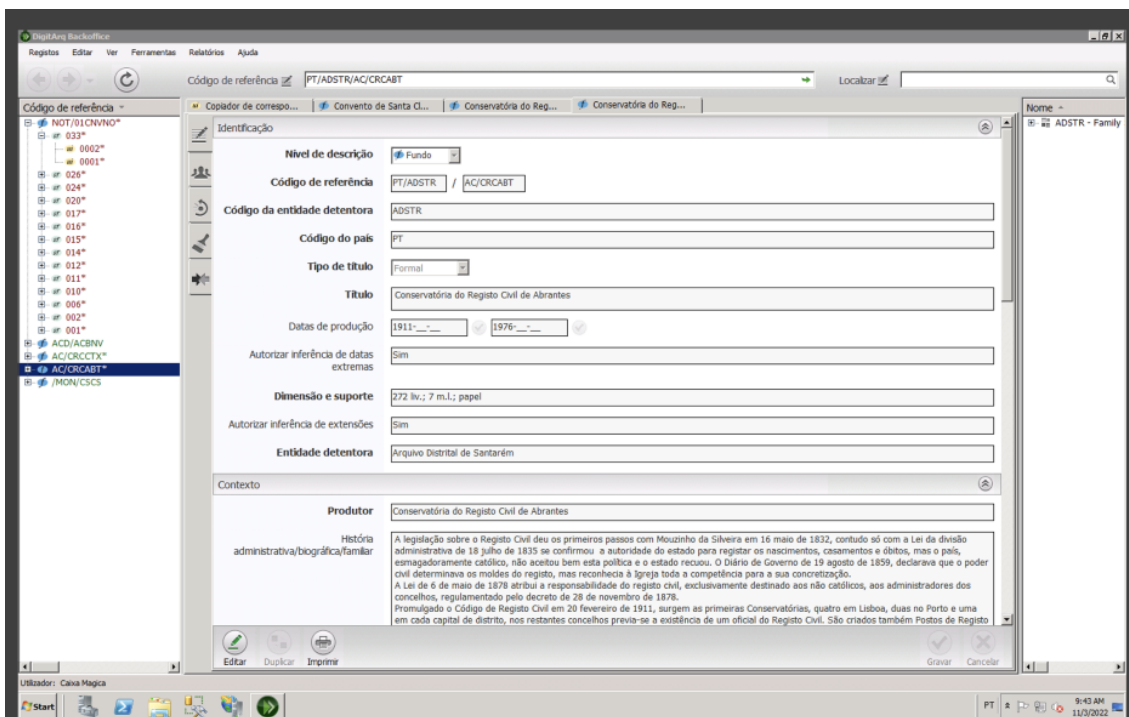
Além disso, estabelecemos um cronograma de reuniões de revisão a cada 15 dias com o cliente, conhecidas como *sprint reviews*. Nestas reuniões, apresentamos as funcionalidades desenvolvidas durante o período de *sprint* e recolhemos o *feedback* do cliente. Isto permitiu-nos ajustar a direção do projeto com base nas necessidades e prioridades do cliente, garantindo uma colaboração próxima e um processo de desenvolvimento ágil, explicado em mais detalhe na secção 4.3.

Para acompanhar as tarefas e mantermo-nos organizados, utilizamos a plataforma Trello (referenciada na secção 4.1). Criámos um quadro no Trello com listas de tarefas, cada uma representando um *sprint*. As tarefas foram divididas em cartões individuais, que foram atribuídos aos membros da equipa. Isto permitiu-nos acompanhar o progresso de cada tarefa, identificar dependências e garantir uma distribuição equilibrada do trabalho.

Para estimar a complexidade das tarefas, utilizamos a técnica de *story points*¹, explicada em detalhe na subsecção 4.1.8. Cada tarefa recebeu uma pontuação de acordo com a sua complexidade, esforço e riscos envolvidos. Esta abordagem ajudou-nos a priorizar as tarefas e alocar recursos de forma eficiente, garantindo que os prazos fossem cumpridos e que a qualidade do trabalho não fosse comprometida.

Esta abordagem metodológica foi fundamental para manter a equipa alinhada, gerir as expectativas do cliente e garantir a entrega de uma plataforma Digitalq robusta, funcional e com uma experiência de utilizador melhorada.

¹<https://asana.com/pt/resources/story-points>



1.1 Registo de descrição

- É possível editar e imprimir qualquer ficheiro.
É possível duplicar qualquer ficheiro **excepto fundos**.
- É possível expandir/esconder os painéis um a um ou todos ao mesmo tempo com duplo clique.
- É possível abrir várias tabs com diferentes documentos. A forma como as tabs são abertas não está clara. Se eu carregar num documento à esquerda ele substitui a tab aberta. Só consigo abrir uma nova tab se for à barra de ferramentas em Registos > Abrir fundo.
- É possível adicionar um novo registo dentro de um documento ao carregar com o lado direito do rato sobre um documento:

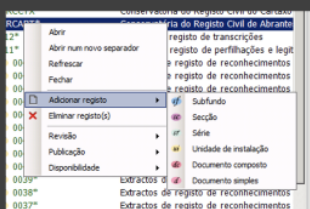


Figura 5.1: Análise do *backoffice* da plataforma DigitArq

Capítulo 6

Trabalho relacionado

O estado da arte, relativamente a websites de descrição arquivística está a evoluir e a ser explorado em vários países. Países como os Estados Unidos da América [3] e o Reino Unido [2] tem atualmente as suas plataformas arquivísticas atualizadas para corresponder às necessidades dos cidadãos nacionais, proporcionando uma experiência de utilização moderna, intuitiva e ainda um *back-end* redesenhado para as buscas serem mais otimizadas.

Portugal segue a norma internacional ISAD(G), o que significa que a nível arquivístico, o modelo de organizar os arquivos é bastante uniformizado, resultando em plataformas algo similares na sua abordagem à disponibilização de documentos a nível nacional. A DGLAB para além de seguir o modelo convencional de representações, vai mais além com a sua abordagem de querer inovar a sua experiência de utilização, bem como o modelo de dados e mudança de paradigma na sua base de dados e consequente migração. Em termos de representação de entidades, consequentemente fundamentais para boas práticas de web semântica, a DGLAB pretende usar no novo Digitarq o modelo CIDOC-CRM, uma forma de representar objetos culturais, que é atualmente usada em organizações como o British Museum [12] e pelo Museu del Prado [6].

6.1 Museu da Pessoa

Outro exemplo que junta CIDOC-CRM e web semântica é o projeto Museu da Pessoa [5], no Brasil. Este é um projeto criado em 1991, por um grupo de historiadores que decidiu construir uma plataforma onde fossem armazenados testemunhos de vida de todas as pessoas, famílias, comunidades ou instituições. O objetivo é que seja retratado todo o tipo de indivíduos, do mais famoso ao mais comum e anónimo. Ao contrário de um museu normal, este representa pessoas, e não objetos físicos em exposição. Foi em 1999 que a Universidade do Minho decidiu começar a colaborar neste projeto com o compromisso de ajudar a enriquecer a coleção de histórias de vida, e ainda, criar novas tecnologias que permitam explorar em detalhe, e de forma inteligente, o conteúdo presente nesta coleção [1]. O processo para a recolha destes testemunhos consiste em quatro passos, sendo eles: uma entrevista gravada com áudio ou vídeo, a transcrição dessa mesma entrevista, a anotação da transcrição na linguagem XML e por fim estes documentos podem ainda ser usados para produzir diferentes tipos de output. A ontologia criada para este projeto

consistiu, numa primeira fase, por criar uma ontologia própria, baseada na extração de conceitos e entidades presentes nas entrevistas às pessoas, seguido de um mapeamento dessa informação para o modelo CIDOC-CRM, com o objetivo de adotar um modelo padrão na área dos museus, e, por fim, foi feita a adição de outros dois vocabulários, o FOAF (*Friend of a Friend*), e DBpedia, devido ao enriquecimento que estes poderiam dar ao modelo previamente desenvolvido. Depois desta descrição da informação, o Museu da Pessoa começou a utilizar RDF para transformar documentos XML, em dados estruturados, em forma de triplos RDF.

O Digitarq assemelha-se com este projeto no sentido em que irá utilizar a ontologia do CIDOC-CRM e o modelo RDFa (uma forma de representar RDF em HTML) para descrever as entidades arquivísticas.

6.2 Europeana

O projeto Europeana [8] é uma iniciativa de destaque na gestão e acesso a recursos culturais e patrimoniais digitais. A Europeana adota uma abordagem abrangente, oferecendo uma plataforma que permite o acesso e a descoberta de uma vasta gama de conteúdos culturais provenientes de instituições em toda a Europa. Esta plataforma exibe material de herança cultural de cerca de 3.700 instituições diferentes. Essa informação chega até à Europeana através de parceiros agregadores que recolhem os dados, verificam-nos e enriquecem-nos com informações, tais como geolocalização ou conectando esses recursos a outras informações ou conjuntos de dados usando pessoas, locais ou tópicos relacionados - ver figura 6.1



Figura 6.1: Processo de agregação de informação utilizado pela Europeana

No que diz respeito às tecnologias utilizadas na Europeana, a plataforma faz uso da framework Nuxt.js para o desenvolvimento do *front-end*. O Nuxt.js é uma estrutura de aplicação web em JavaScript baseada na popular framework Vue.js. Esta simplifica a criação de aplicações Vue.js com renderização do lado do servidor (SSR), pré-renderização estática e gestão de rotas. O Nuxt.js facilita a criação de interfaces de utilizador reativas, eficientes e escaláveis, fornecendo uma base

sólida para a construção de aplicações web modernas.

No *back-end*, a Europeana adota uma arquitetura robusta para a gestão e entrega de dados culturais digitais, envolvendo tecnologias como Java, Python ou Ruby, juntamente com frameworks como Spring ou Django, para processamento, armazenamento e recuperação eficiente de grandes volumes de dados.

Quanto ao modelo de dados, a Europeana utiliza o Europeana Data Model (EDM), um modelo conceitual projetado para acomodar diversos tipos de recursos culturais, como texto, imagens, áudio e vídeo. O EDM adota princípios de dados ligados, como o uso de RDF e OWL (*Web Ontology Language*), permitindo a criação de conexões semânticas e relações entre os recursos culturais. Além disso, a Europeana incorpora o CIDOC-CRM como parte do seu modelo de dados, oferecendo uma estrutura comum para a descrição de objetos culturais, eventos, atores e suas relações.

Essas tecnologias e modelos de dados utilizados pela Europeana apresentam semelhanças com o projeto Digitarq. Ambos os projetos têm como objetivo fornecer uma plataforma de gestão arquivística que permite a descrição, acesso e preservação de recursos culturais de forma padronizada e interoperável. O Digitarq utiliza a framework Next.js para desenvolver o *front-end*, permitindo a criação de interfaces de utilizador reativas e eficientes, semelhante ao que acontece na Europeana. Além disso, o Digitarq reconhece a importância do modelo de dados CIDOC-CRM para garantir uma representação estruturada e semântica das informações culturais.

Com base nessas semelhanças, podemos afirmar que o Digitarq está alinhado com abordagens modernas e globais de gestão arquivística, seguindo os passos da Europeana na busca por um modelo global e moderno. Através da adoção de tecnologias avançadas e modelos de dados interoperáveis, o Digitarq está posicionado para se tornar uma plataforma abrangente e escalável, permitindo uma gestão eficiente e um acesso facilitado aos recursos arquivísticos.

Baseado nesta análise, surge então uma comparação natural entre estas duas plataformas. Por um lado temos o Digitarq, uma plataforma nacional que ambiciona modernizar a forma como trata os dados e os disponibiliza para os utilizadores. No outro lado do espectro temos a Europeana, uma plataforma que utiliza os dados disponibilizados por instituições ao redor do mundo, para compor a sua interface. Ambas as plataformas têm em comum a modernização nas suas abordagens, mas diferem no seu propósito final. Mais do que divergir, estas plataformas podem ser complementares, na medida em que o Digitarq, que suporta já o uso do protocolo OAI-PMH, mencionado na secção 4.4, pode tanto beneficiar da sua nova interface redesenhada, como, futuramente, e se for essa a intenção, contribuir para a vasta rede de parceiros da Europeana, desta maneira, concentrando toda a informação numa só plataforma.

Capítulo 7

Arquitetura

A arquitetura do projeto, conforme ilustrado na figura 7.1, foi projetada para permitir uma escalabilidade fácil, proporcionando uma solução baseada em microserviços, que, por sua vez, estarão inseridos em contentores Docker, mencionado na subsecção 4.1.9. A escalabilidade é oferecida pelo acréscimo ou decréscimo de capacidade, ou seja, à medida que a carga aumenta, o número de contentores aumenta de acordo, e, caso contrário, estes são desativados. Existirá, portanto, um sistema de orquestração, utilizando Kubernetes, que irá comandar estas alterações na capacidade do sistema e que permitirá acompanhar o crescimento da aplicação. Ao momento de escrita deste trabalho, com o sistema ainda em fase de desenvolvimento, o sistema de orquestração está inativo, estando em vigor um número fixo de contentores.

Servidores Nginx (ver subsecção 4.1.9) estão na base desta estrutura, sendo assim uma solução *self-hosted*. Esta solução proporciona um maior controlo, uma melhor gestão da informação e um controlo total dos backups. No entanto, apresenta um maior esforço na configuração e manutenção dos sistemas.

Existem contentores específicos para o *front-end*, que utilizam o React para servir a aplicação; para o *back-end*, que é baseado em Node.js; para uma base de dados relacional MySQL, onde está contida informação relativa aos utilizadores e todo o tipo de informação que não seja proveniente dos arquivos; uma base de dados de grafos Neo4J, que contém dados relacionados com os arquivos migrados para este novo paradigma; e ainda uma base de dados Elasticsearch, que possui a mesma informação que a base de dados de grafos, mas que é usada para obter pesquisas de forma mais eficaz e inteligente.

Esta arquitetura bem definida e a escolha cuidadosa das tecnologias e bases de dados garantem que a plataforma Digitarq seja robusta, escalável e capaz de lidar com o crescimento e a crescente demanda de utilizadores e dados. A combinação de microserviços, contentores Docker e a utilização de tecnologias modernas e eficientes, como Next.js, Node.js, MySQL, Neo4J e Elasticsearch, cria uma base sólida para a construção e o desenvolvimento da plataforma Digitarq.

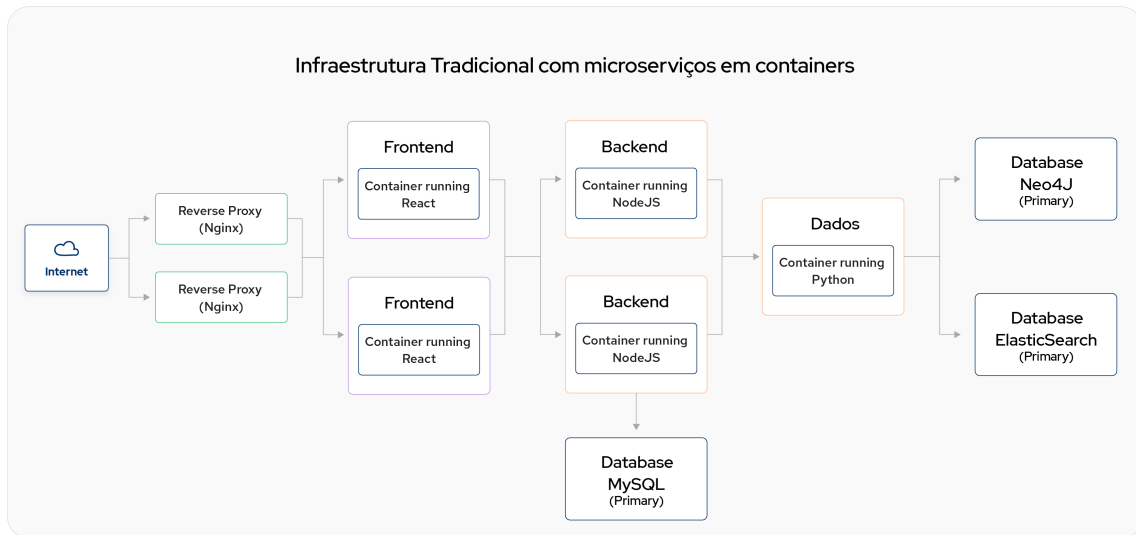


Figura 7.1: Arquitetura do Projeto

Capítulo 8

Trabalho realizado

Nesta secção será relatado todo o trabalho completado até à conclusão deste projeto de tese, bem como as dificuldades apresentadas no decorrer das tarefas e alguns pormenores técnicos da implementação da aplicação.

8.1 UI/UX aplicado ao projeto

Para este projeto foi planeado existirem interfaces com uma boa experiência de utilizador e, para isso, foram realizadas duas sessões de testes de usabilidade no início do projeto, e mais uma sessão no decorrer do mesmo. Estes testes com utilizadores foram realizados como parte dos objetivos específicos da tese.

8.1.1 Testes iniciais

Os testes iniciais do projeto foram realizados tanto com utilizadores da plataforma atual do Digitarq, como com utilizadores com zero experiência na plataforma e em quaisquer soluções de gestão arquivística. A questão de pesquisa para este estudo foi: "Será o design do novo Digitarq acessível e usável para os utilizadores externos do Digitarq, sendo que estes compreendem uma faixa etária acima dos 40 anos?". Esta questão surgiu devido ao elevado número de pessoas com idade mais avançada, a utilizar o sistema atual do Digitarq, por se tratar de uma ferramenta desatualizada, mas utilizada há largos anos pelos investigadores. Aqui, a variável independente sendo as tarefas a ser realizadas pelos utilizadores, e a variável dependente, o número de tarefas que um utilizador conseguiu realizar com sucesso. Para além deste dado quantitativo numa escala de três possíveis valores sendo eles: conseguiu realizar a tarefa, conseguiu realizar com ajuda, ou, não conseguiu, recolhemos ainda dados qualitativos consistindo em opiniões ou informações que os participantes forneciam no decorrer do teste, com base na sua experiência prévia com a plataforma antiga. Este estudo consistiu num estudo *within-subjects* (estudo em que todos os participantes de um grupo avaliam todas as condições da experiência), separado entre dois dias diferentes devido à necessidade de se melhorar o design com base em feedback prévio. Este teste de usabilidade foi apenas ao design e respetiva prototipagem elaborada pela equipa de design da CMS.

Numa primeira instância foram feitos quatro testes piloto para averiguar se a estrutura do teste

estava adequada e se eram detetados erros ou necessários alguns ajustes. Após passar os testes pilotos com sucesso, foi realizado o teste com utilizadores na sede da DGLAB e contou com a participação de 20 indivíduos no primeiro dia, e 8 no segundo, com uma duração de teste de aproximadamente 30 minutos, sendo que existiram sempre dois testes a decorrer em simultâneo. Os participantes tinham idades compreendidas entre os 40 e 60 anos para corresponder à necessidade de responder à nossa questão de pesquisa e foram escolhidos pela DGLAB, sendo trabalhadores internos da instituição, deste modo, não sendo recompensados com nenhum benefício. Começaram-se os testes por disponibilizar aos participantes uma página com informações relativa à sua participação, nomeadamente que a sua identidade seria protegida e que iria ser feita a gravação de áudio e cliques no ecrã, para uma melhor análise dos resultados obtidos. Depois da assinatura recolhida, passámos a dar uma breve introdução ao estudo, explicando o que os utilizadores iriam observar e como iria decorrer o processo. Após esta introdução, apresentaram-se algumas tarefas relacionadas com a utilização da interface: algumas meramente visuais e de feedback auditivo, outras mais práticas, implicando a utilização do rato do computador para navegar. Ao longo das tarefas foi-se registrando a performance dos participantes de modo quantitativo, assinalando se tinham ou não conseguido completar a tarefa, e ainda, qualitativo, descrevendo quais as dificuldades enfrentadas, o pensamento que os utilizadores realizavam ao longo das tarefas e possíveis notas.

Na generalidade dos casos os testes correram de acordo com o esperado e nos tempos previstos. O design foi bem recebido pelas pessoas, embora os utilizadores com experiência na plataforma antiga (cerca de 80%), tenham apresentado algumas dificuldades, prendendo-se com o facto de estarem habituados à mesma interface há um tempo considerável. Foi ainda observada a possibilidade de existir uma nova funcionalidade que não estava no plano de trabalhos e que foi fortemente recebida pelos utilizadores, aquando da formulação da ideia em testes posteriores. Esta funcionalidade prende-se com o facto de os utilizadores poderem guardar documentos em diversas pastas, ao invés de guardar tudo numa só lista. Esta funcionalidade traz aos investigadores a possibilidade de melhor organizar o seu trabalho, para além de ser já um sistema bastante utilizado noutras plataformas como as de gestão de ficheiros.

Uma alteração importante que se realizou após a primeira sessão dos testes foi no ecrã dos resultados da pesquisa, onde se removeu o painel lateral dos filtros, representado na Figura 8.1, para um modal, simplificando a informação. Podemos ver esta alteração nas figuras 8.2 e 8.3. Foram ainda adicionadas labels em alguns botões que apenas apresentavam ícones e uma descrição da hierarquia de documentos, na página de detalhes do mesmo. Por fim, o contraste foi aumentado em toda a interface. Após a segunda iteração dos testes, removeu-se os filtros da página inicial, por não serem relevantes para os utilizadores. Associámos essa mudança ao que acontece no maior motor de busca, o Google, que na sua página inicial apenas apresenta um campo de pesquisa, apresentando apenas mais tarde os filtros, o que simplifica o processo inicial de procura. Aqui o objetivo é que, apenas a partir dos termos de pesquisa, os nossos resultados sejam relevantes o suficiente para minimizar o uso de pesquisas avançadas.

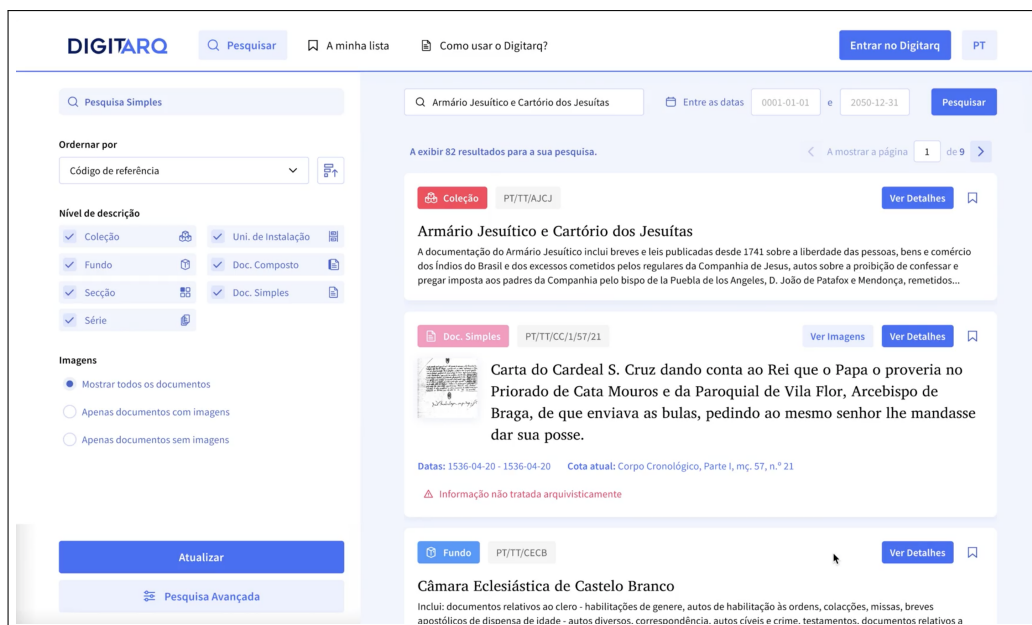


Figura 8.1: Ecrã dos resultados de uma pesquisa, na primeira fase de iterações dos testes de usabilidade

No primeiro dia de testes obtivemos uma média de 23.5 respostas em que o participante conseguiu responder, ou indicar, corretamente, de um total de 32 perguntas, resultando em aproximadamente 73% de precisão nas respostas. No segundo dia, a média de respostas corretas foi de 28.1, desta vez em 33 perguntas, obtendo-se um valor aproximado de 85% de precisão.

Concluimos que estes testes foram bastante produtivos do ponto de vista do melhoramento da interface, contando com uma subida de doze por cento na média de precisão de respostas certas de um teste para o outro. Embora esta subida seja considerável, não dá garantias que não tenha sido influenciada por fatores externos como por exemplo, um dos participantes do primeiro dia informar um segundo participante do segundo dia acerca do conteúdo dos testes, ficando este mais ciente da experiência, e consequentemente ter um melhor desempenho.

8.1.2 Testes Conservação e Restauo

A área das intervenções abrange os diagnósticos, bem como as ações de Conservação e Restauo aplicadas a um documento. Esta parte do site é exclusiva a utilizadores internos estando por isso situada no *backoffice* da aplicação.

Antes da elaboração prática desta área tiveram de ser realizados testes de usabilidade devido à sua complexidade e ao facto de a DGLAB não estar habituada a utilizar a sua antiga plataforma para documentar este tipo de relatórios. Ao invés do *Digitalarq backoffice*, os trabalhadores utilizam, na grande maioria dos casos, papel e caneta, devido à má experiência de utilização e complexidade dos formulários. Após a equipa de design ter desenhado um mockup do que poderia ser a área de Conservação e Restauo, iniciámos os testes com utilizadores para compreender se esta nova interface iria ser intuitiva o suficiente para começar realmente a ser utilizada, desta forma,



Figura 8.2: Ecrã dos resultados de uma pesquisa

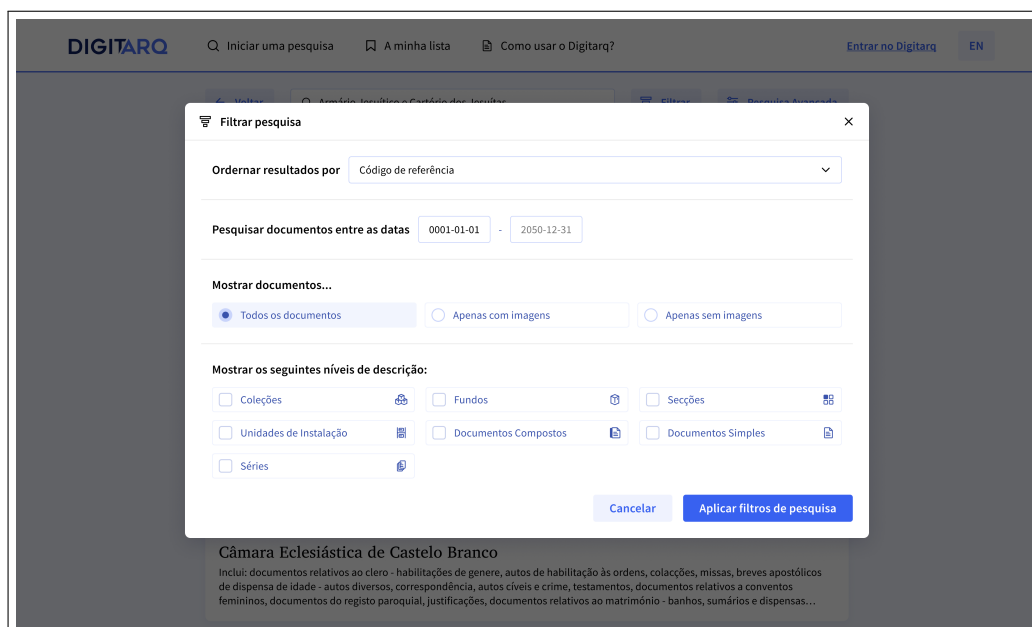


Figura 8.3: Ecrã dos resultados de uma pesquisa, com o modal dos filtros de pesquisa aberto

permitindo que os dados ficassem corretamente armazenados e estruturados.

Para realizar os testes utilizámos uma plataforma online - Maze (ver subsecção 4.1.5), que permitiu aos participantes terem uma experiência mais fidedigna nos testes, uma vez não dependem de sermos nós a dar as instruções. Nesta plataforma, as descrições das tarefas aparecem detalhadas no ecrã, e estão acessíveis sempre que o utilizador as quiser rever. Para além disto, enquanto os utilizadores realizam o teste, o Maze regista os cliques no ecrã, a gravação do áudio, o tempo por tarefa e mais alguns dados estatísticos para futura análise. Importante também destacar que o foco principal desta plataforma é a possibilidade de realização de testes à distância, contudo, achámos por bem realizar na mesma os testes presencialmente, devido à complexidade da interface apresentada e por estarmos a lidar com utilizadores menos experientes com plataformas tecnológicas.

Estes testes de usabilidade foram realizados numa sessão apenas e tiveram as mesmas características (questão de pesquisa e variáveis) que a prévia sessão que foi realizada no início do projeto. Desta vez surgiu apenas uma mudança na variável dependente: número de tarefas realizadas com sucesso, em que os valores possíveis para determinar se uma tarefa foi concluída são: sucesso direto, sucesso indireto e não completou. Esta distinção deve-se ao facto de a própria plataforma registar assim os valores, através do caminho do utilizador até completar, ou não, cada tarefa. Neste estudo contámos com a presença de 11 participantes, sendo que 10 deles eram do sexo feminino e 1 do sexo masculino. A idade média dos participantes foi de aproximadamente 39 anos, tendo eles idades compreendidas entre os 23 e 62 anos. De salientar que tanto o número de participantes, como as suas idades ou sexo, não foram fatores escolhidos pela equipa de UI/UX, mas sim fatores aleatórios que ocorreram derivados dos participantes que a DGLAB havia disponibilizado. Registámos ainda o tempo que estiveram a trabalhar como técnicos da Conservação e Restauro, e consequentemente expostos à utilização do antigo Digitarq e obtivemos uma média de 17 anos, uma mediana de 22, e um desvio padrão de 12.2, sendo que não temos os dados de um dos participantes (P2). Podemos ver estes valores na figura 8.4.

A compreensão estatística da influência da idade e da experiência dos participantes na adaptabilidade a uma nova interface proposta é um aspecto interessante a ser explorado. Ao analisar esses dois fatores, é possível ter percepção sobre como diferentes faixas etárias e níveis de experiência podem afetar a capacidade dos utilizadores em se adaptar a essa interface. Através da criação de um gráfico de dispersão 8.5, pudemos observar uma relação significativa entre essas variáveis. A experiência dos utilizadores mostrou-se como o fator mais influente, com um coeficiente de determinação (R^2) - métrica estatística que fornece uma medida de quão bem um modelo estatístico se ajusta aos dados. Ele indica a proporção da variabilidade da variável dependente que pode ser explicada pelas variáveis independentes do modelo, variando de 0 a 1, onde 0 indica que o modelo não explica nenhuma variabilidade e 1 indica que o modelo explica toda a variabilidade - de 0.553. Esse valor indica uma correlação moderadamente forte entre a experiência e a capacidade de adaptação à interface proposta. Por outro lado, a idade dos participantes também apresentou um valor de 0.147, o que sugere uma influência menor, porém ainda relevante, na facilidade de uso da interface.

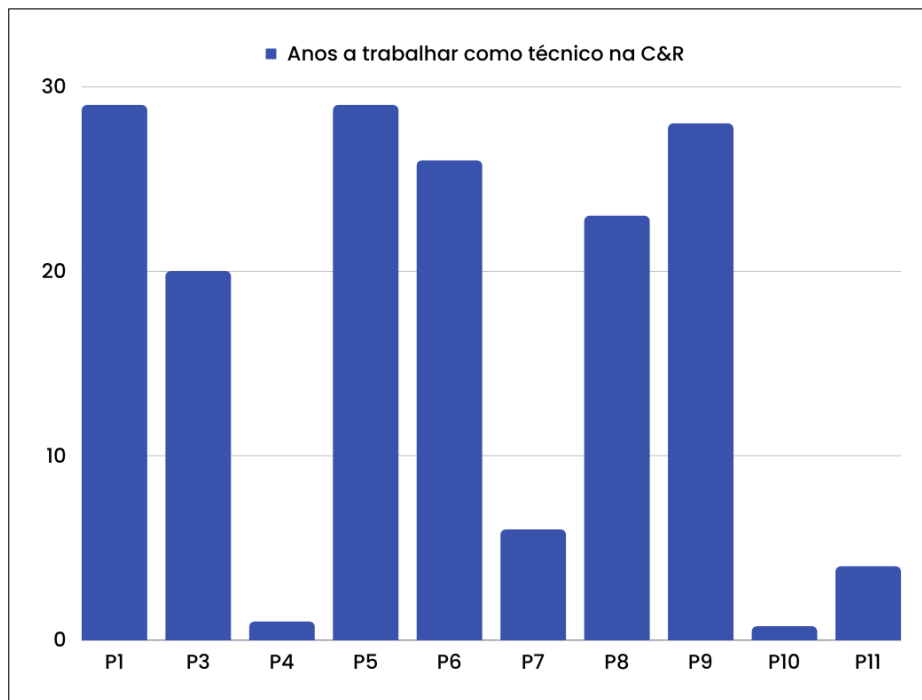


Figura 8.4: Anos de experiência dos participantes no estudo

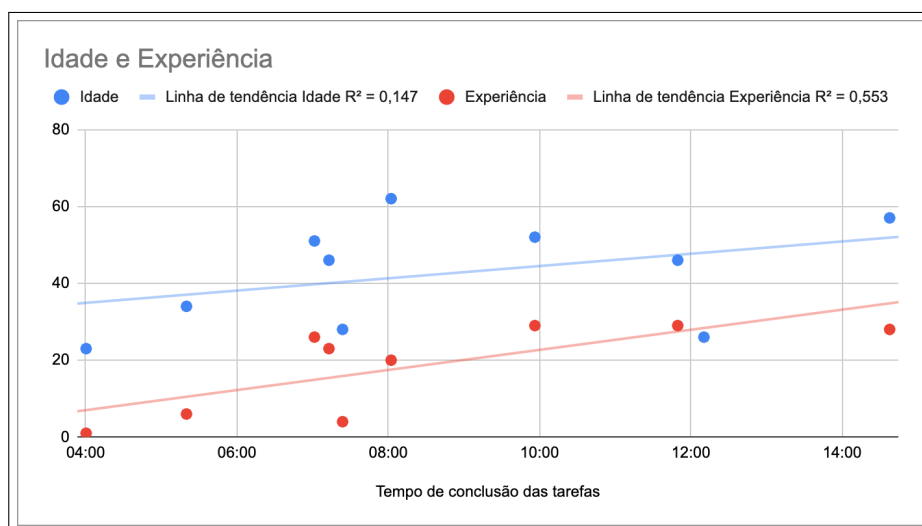


Figura 8.5: Correlação entre a idade, a experiência e o tempo total de conclusão das tarefas

Além destas análises estatísticas que realizámos, conseguimos obter outros dados interessantes através da plataforma Maze, nomeadamente as taxas de sucesso de cada utilizador em dada tarefa, a taxa de erro de cliques, onde no ecrã é que sucederam os cliques e os caminhos que o utilizador usou para concluir cada tarefa. Na figura 8.6 podemos observar que na primeira tarefa, em que o objetivo era criar uma ação de recuperação num dado documento, para abrir esse mesmo documento houve um utilizador que escolheu o caminho menos provável, contudo, um caminho possível. Este tipo de análises são bastante relevantes para perceber onde a interface pode ser melhorada para ir de encontro a uma experiência cada vez mais intuitiva. Já na figura 8.7 podemos ver exatamente onde ocorreram os cliques no primeiro ecrã da primeira tarefa. Podemos ainda observar o mapa de calor da última tarefa, em que era pedido ao utilizador que abrisse os detalhes de uma ação em específico 8.8. Desta vez constata-se que houve mais dificuldade por parte dos utilizadores em clicar corretamente na opção certa, embora não seja uma percentagem de cliques errados suficientemente elevada que justificasse uma alteração na interface.

Ao longo do relatório do Maze podemos observar valores de *misclick* bastante elevados, inclusive alguns valores a 100 por cento, mesmo quando a tarefa foi concluída com sucesso. Um exemplo disso é na figura 8.7, que embora vejamos uma nuvem de volume de cliques muito grande no botão principal para avançar a tarefa, a plataforma indica que para este ecrã a taxa de cliques errados foi 100 por cento. Não conseguimos determinar se estes valores se devem a algum erro ao montar o protótipo, mas, por esta razão, este fator não foi levado muito em conta aquando da análise dos dados. No geral, os testes correram conforme o esperado, com os utilizadores a terem um feedback bastante positivo sobre esta área e por essa razão não foi realizada nenhuma alteração ao design inicial. Como notas finais, é importante realçar a relevância desta plataforma na realização dos testes, que é capaz de fornecer dados que, doutra maneira, seriam bastante complicados de obter, e atentar também para a duração e complexidade das tarefas. Um dos aspetos sentidos durante as entrevistas foi que as tarefas estavam algo densas e com algum grau de dificuldade (como por exemplo dizer ao utilizador para abrir um determinado fundo, sendo que o título desse mesmo fundo era longo), o que frequentemente distraía o participante devido a ter de voltar a ler as instruções a meio da tarefa.

8.2 Frontoffice

Nesta secção vão ser descritos os vários ecrãs e páginas que compõem a aplicação do lado do utilizador externo que visita o Digitarq, bem como explicar algumas decisões técnicas que foram tomadas.

8.2.1 Inicialização

Quanto ao desenvolvimento inicial, após os testes de usabilidade, foi iniciada a componente *frontend* do projeto, utilizando alguns componentes do projeto CRAV, desenvolvido pela Caixa Mágica Software e mencionado no capítulo 4. A componente de autenticação foi uma parte que foi reutilizada para o novo projeto, sendo um login único entre estas duas plataformas. Foi criado um novo

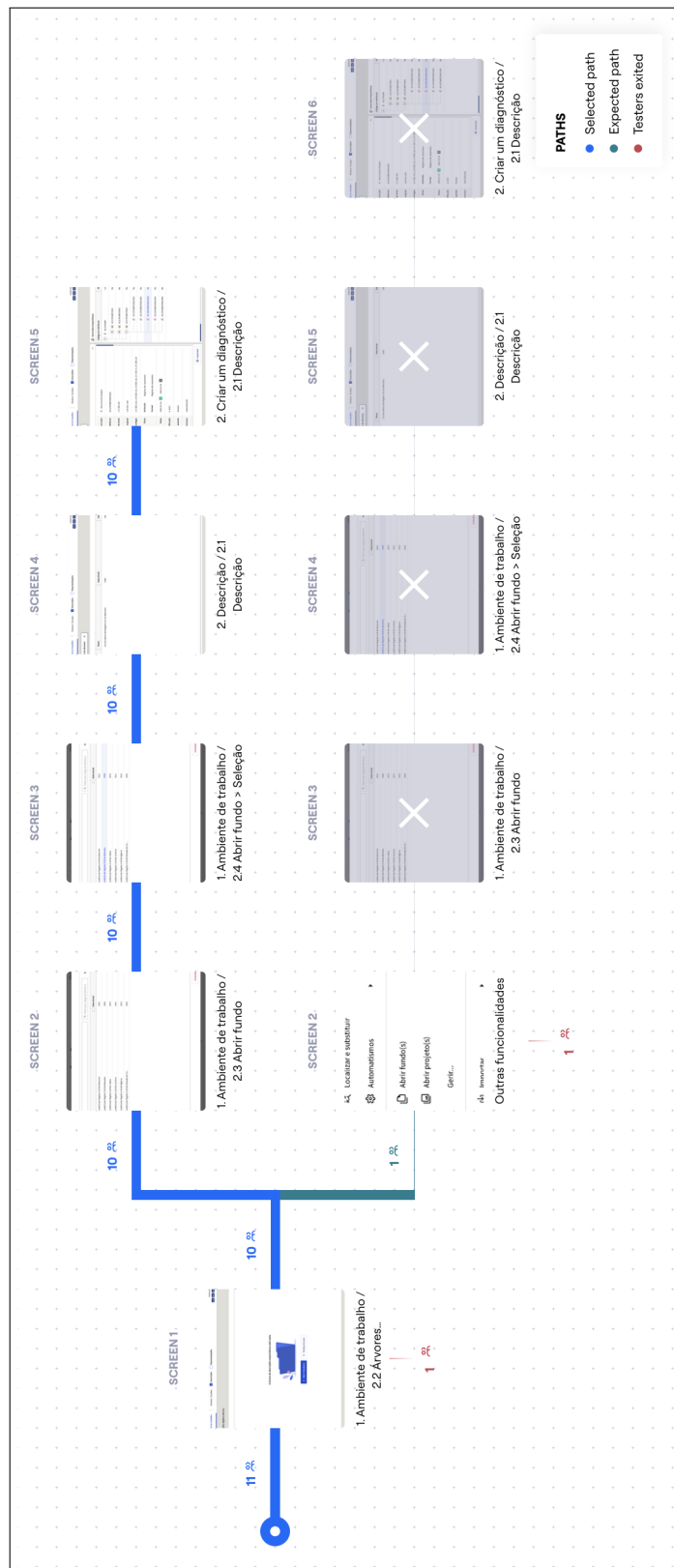


Figura 8.6: Caminho dos utilizadores para concluir a tarefa: criar uma ação de recuperação num dado documento



Figura 8.8: Mapa de calor da tarefa de ver detalhes de uma ação

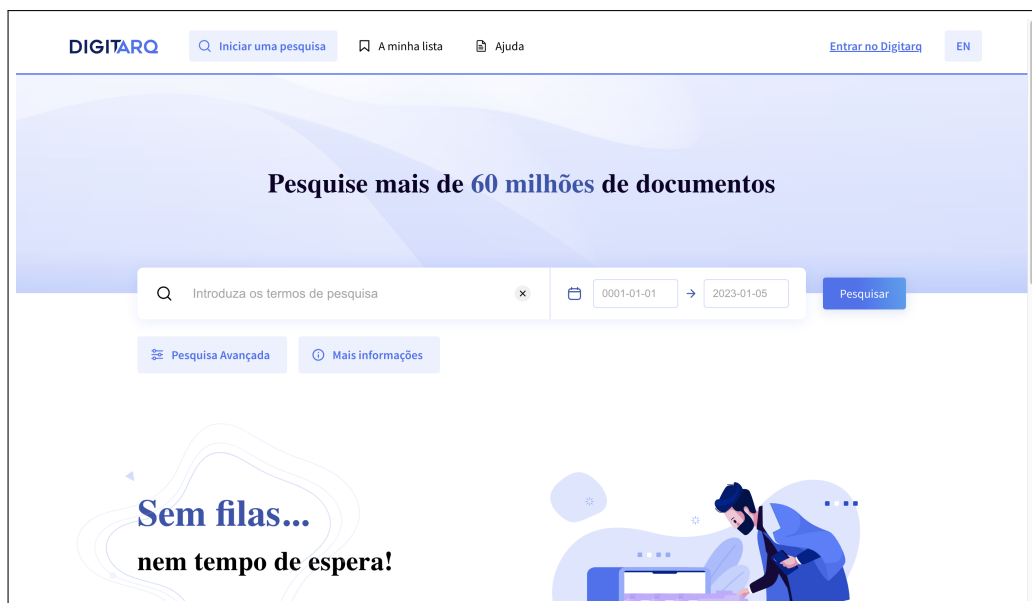


Figura 8.9: Página inicial, implementada no Digitarq

projeto no *gitlab* com a inicialização do projeto em Next.js. A componente do *front-office* foi a primeira a ter início e conta com dez páginas diferentes, sendo elas: a página principal; o login; a página da pesquisa simples¹; pesquisa avançada; detalhes do documento¹; detalhes da entidade; visualizador de imagens¹; arquitetura em árvore de um documento¹; página de documentos guardados¹ e secção de ajuda. Numa primeira abordagem começou-se por desenvolver muitos dos componentes que iriam ser reutilizados ao longo do site, utilizando componentes de React e CSS.

8.2.2 Página inicial e login

Após a inicialização do projeto, a construção de alguns componentes base seguindo a estrutura atômica de ficheiros, e da implementação de propriedades do design system, seguiu-se o desenvolvimento do login e da página inicial, que contém inputs para uma pesquisa simples e alguma informação e grafismos. A página inicial consiste essencialmente numa caixa de pesquisa que permite aos utilizadores começarem a pesquisar por termos e datas, os documentos que estes procuram. Esta pesquisa comunica com a API de dados que está responsável por fazer queries à base de dados Elastic Search. A partir desta página é ainda possível navegar até à pesquisa avançada. Ver Figura 8.9.

8.2.3 Pesquisa Simples

A página da pesquisa simples é onde são apresentados os resultados da pesquisa realizada na página inicial. Estes resultados consistem em cartões com informação básica sobre os documentos, tal como o nível de descrição, o código de referência, o título, as datas, quotas e ainda uma miniatura de uma imagem do documento. Para além de serem mostrados os resultados, é possível

¹Páginas realizadas no âmbito deste trabalho de tese

ainda filtrar os mesmos, realizar nova pesquisa, saltar para a pesquisa avançada e ainda navegar entre a paginação dos resultados. Para além de visualizar estas informações para cada documento, é ainda possível, caso esteja registado na plataforma, guardá-los para uma consulta posterior. Esta funcionalidade da lista de itens guardados irá ser explorada mais à frente no relatório.

8.2.4 Detalhes do Documento

Nesta página (ver figura 8.10) é possível consultar informações mais detalhadas sobre um documento, é possível saltar para um determinado campo/detalhe do documento, guardar o documento numa pasta do utilizador, aceder a um popup de ajuda relativamente ao ecrã atual, saltar para o visualizador das imagens e ainda visualizar uma estrutura em árvore simplificada, contendo os documentos, ou as estruturas arquivísticas no nível diretamente abaixo e acima. Para além dos detalhes, é possível ainda ver as entidades, como os atores, eventos e lugares que estão associados de alguma maneira a esse mesmo documento.

Foi nesta página que começaram a surgir as primeiras questões relacionadas com a estrutura em que chegavam os dados do *back-end* para serem apresentados no *front-end*. Aqui, era necessário que chegasse informação relativa ao documento em si que estamos a ver em detalhe, mas também informação sobre a estrutura hierárquica do documento. Apesar desta informação ser muito densa, considerámos que nem toda era relevante para ser apresentada: apenas era necessário o nível diretamente abaixo na árvore do documento. A solução encontrada foi ter um parâmetro na rota da API do *back-end* que expressasse quantos níveis são necessários. Para além de passar apenas um nível, o que reduz bastante a carga de informação, decidimos também passar apenas os parâmetros imprescindíveis a ser mostrados no *front-end*, sobre os documentos da estrutura hierárquica do documento atual.

Esta comunicação entre a equipa do *front-end* e *back-end*, e, mais à frente, com a equipa de dados, é essencial no decorrer do projeto. Nesta fase surgiu também a necessidade de se procurar por bibliotecas de código aberto que permitissem a visualização das imagens, bem como ampliar, ver todas as miniaturas e que suportasse ficheiros nos formatos tiff e jpeg. A página de visualização de imagens será mostrada mais à frente.

8.2.5 Visualizador de Imagens

No âmbito de conseguir mostrar as imagens do documento, e, mais tarde, imagens no backoffice, foi realizada uma pesquisa de bibliotecas open-source, mencionada no capítulo 4. Esta página (ver figura 8.11) é assim composta por uma secção com miniaturas e uma secção principal onde é mostrada a imagem selecionada bem como os comandos de rodar, fazer zoom, navegar entre as restantes imagens e descarregar. Este visualizador é ainda reaproveitado na secção das representações digitais no backoffice.

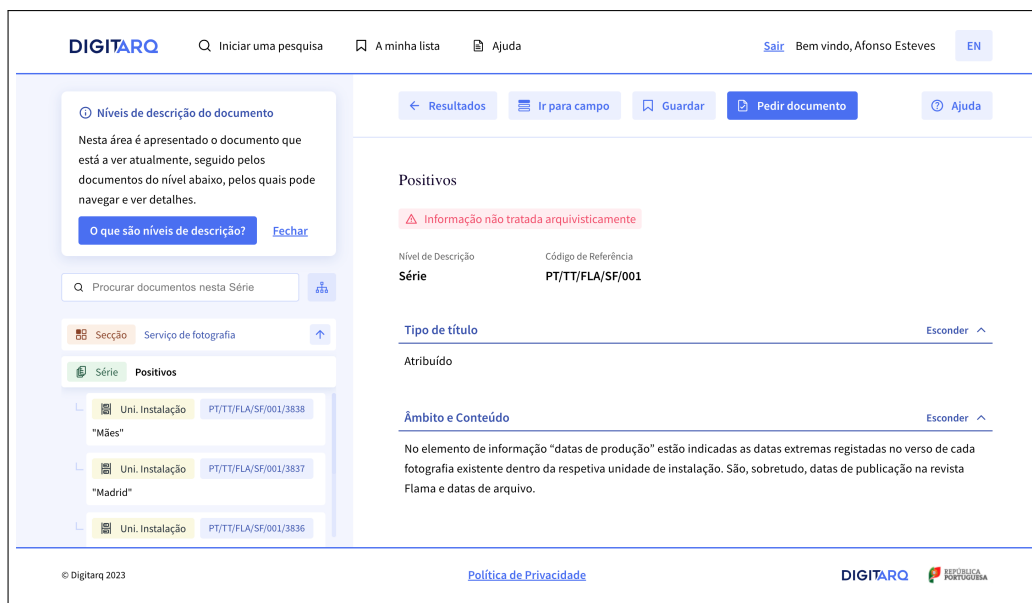


Figura 8.10: Página dos detalhes de um documento

8.2.6 Detalhes das entidades

A página dos detalhes de uma entidade é acessível a partir dos detalhes de um documento, ao clicar num cartão dessa mesma entidade. Aqui, semelhante ao que acontece no documento, são mostradas todas as informações de um ator, lugar ou evento, e ainda as suas associações às restantes entidades. Desta maneira, investigadores terão o seu trabalho bastante mais facilitado ao poderem navegar numa rede de nós totalmente interligada e com referências reais, de uma forma intuitiva através da interface.

8.2.7 Árvore hierárquica do Documento

A página da árvore hierárquica de um documento desempenha um papel fundamental ao apresentar ao utilizador todos os documentos dentro da sua hierarquia, oferecendo também opções de filtro para facilitar a sua visualização. Essa funcionalidade é especialmente importante considerando que os arquivos podem possuir hierarquias com centenas de documentos.

A página da árvore hierárquica tem como objetivo fornecer uma visualização clara e intuitiva da estrutura hierárquica dos documentos. Os documentos são organizados em forma de árvore, onde cada nó representa um documento e as conexões hierárquicas são estabelecidas entre eles. Isso permite que os utilizadores compreendam facilmente a relação entre os documentos e naveguem pela hierarquia de forma mais eficiente.

Além da representação visual, a página da árvore hierárquica oferece recursos adicionais para facilitar a visualização e interação com os documentos. Isso inclui a inclusão de filtros, que permitem aos utilizadores aplicar critérios específicos para limitar a exibição dos documentos. Esses filtros podem ser baseados em atributos dos documentos, como data, tipo de documento, autor, entre outros. Com isso, os utilizadores podem personalizar a visualização da árvore hierárquica de

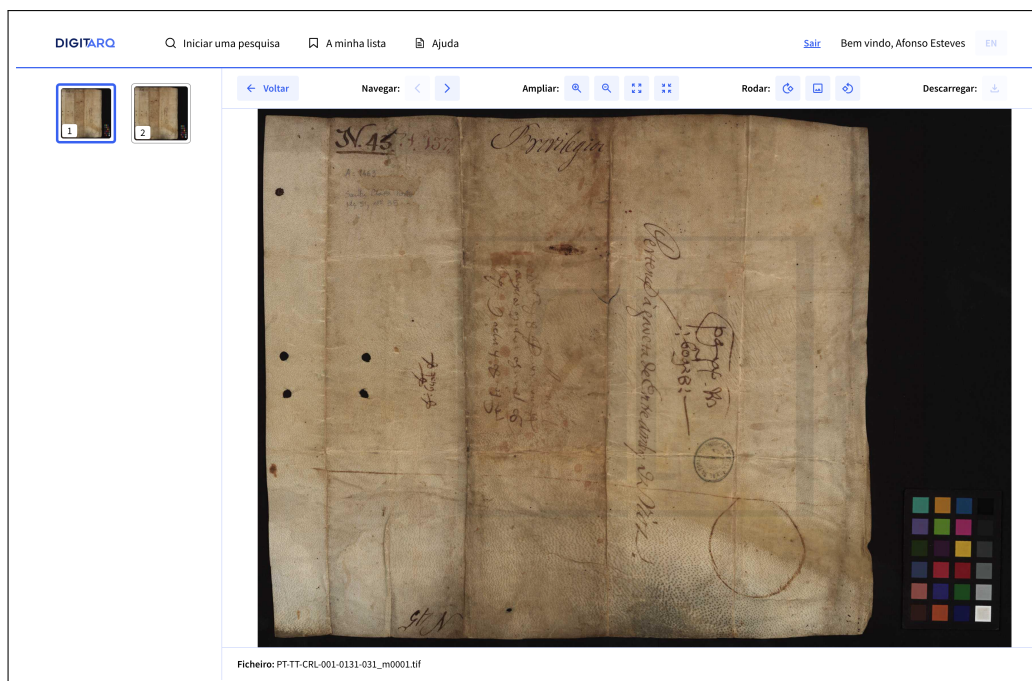


Figura 8.11: Página de visualização de imagens de um documento, no *frontoffice*

acordo com as suas necessidades e preferências, facilitando a localização dos documentos desejados. A implementação desta mesma árvore tornar-se-ia útil mais tarde ao desenvolver as árvores presentes em grande parte do *backoffice*.

8.2.8 Estrutura de pastas

A página da lista de documentos guardados pelo utilizador é uma página a que apenas tem acesso os utilizadores autenticados no Digitarq. Aqui tiveram de ser desenvolvidos novos componentes para as pastas e foi reutilizado um grande fragmento da página da pesquisa simples, sendo esta a lista dos resultados. Nesta página, um utilizador pode ver as suas pastas previamente criadas, criar novas, editar e ainda eliminar pastas, bem como remover documentos das suas pastas. De realçar que quando se navega até esta página, o utilizador fica automaticamente a visualizar a primeira pasta da lista, obrigando sempre a ter no mínimo uma criada para conseguir guardar documentos.

Foi necessário também criar novas rotas no *back-end* que correspondessem aos requisitos desta página. Para além de navegar entre as pastas e os documentos dentro delas, é possível ainda procurar por pastas, procurar por documentos dentro de uma pasta, filtrar a lista dos documentos e imprimir/descarregar a lista. Esta componente de imprimir/descarregar não foi desenvolvida de momento por estar ainda em debate qual das opções é a ideal. Depois de concluída esta página (ver Figura 8.12) e feita a integração com as rotas do *back-end*, foi necessário modificar alguns aspetos na página de pesquisa simples e detalhes do documento, uma vez que, nestas páginas, são incluídos componentes do sistema de pastas. Para ambos os casos foi necessário comunicar com a equipa de *back-end* que seria necessário incluir o token de autenticação nas rotas dessas páginas, um aspeto que não tinha sido pensado previamente, devido ao facto de a pesquisa de

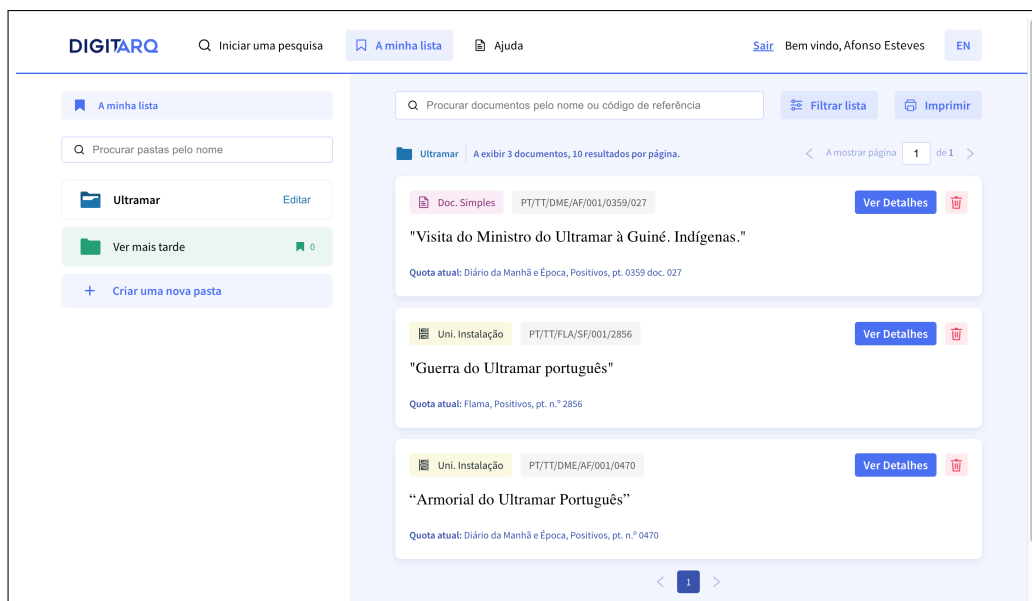


Figura 8.12: Página dos documentos guardados por pastas

documentos ser livre de ser feita por um utilizador não autenticado. Assim, o *back-end* tem de fazer a validação se foi enviado, ou não, um token no pedido. Caso não exista um token, devolve sempre os documentos como não estando guardados em nenhuma lista, pedindo que seja feito o login, na tentativa de guardar. Caso exista um token, será feita a verificação do estado de um documento nas pastas do utilizador, sendo que um documento só pode estar guardado numa pasta. Na página da pesquisa simples para além do estado, é também possível guardar um documento e eliminá-lo da lista do utilizador. Na página dos detalhes, para além destas componentes, é possível ainda criar pastas.

8.3 Backoffice

O *backoffice* é parte essencial da nova plataforma Digitarq, permitindo aos utilizadores internos da DGLAB gerir as informações relativas aos documentos como a descrição dos documentos, as associações que estes têm (Atores, Lugares, Eventos, Documentos), as intervenções de conservação e restauro, as representações digitais e ainda gerir utilizadores da plataforma.

Previamente esta plataforma era uma aplicação *desktop*, mas agora toma a forma de uma aplicação web, juntamente com o *frontoffice*. Esta mudança trouxe algumas dificuldades, principalmente à equipa de UI/UX que teve a difícil tarefa de redesenhar uma plataforma que se apoia muito no uso de tabs, menus e espaço dividido no ecrã, para uma solução web. Ao nível da aplicação surgiram também algumas dificuldades como por exemplo a gestão da informação aberta em cada sessão. Estas dificuldades tiveram de ser abordadas caso a caso e tendo em mente uma experiência similar a uma versão *desktop*. No caso da gestão da informação, resolvemos o problema armazenando a informação relativa a documentos, entidades, ou outro tipo de objetos, na base de dados relacional. Para além de armazenar uma referência para cada um desses objetos,

que na aplicação consideramos as *tabs* (exemplificado na figura 8.13), guardamos ainda o estado da árvore de descrições ou representações digitais, e ainda em que menu é que se encontra o utilizador. Para este efeito, sempre que é feita alguma alteração que envolva um destes três cenários, é feita uma chamada à API do *back-end* de modo a armazenar o estado mais recente do *backoffice* do utilizador.

Ao longo desta secção vão ser abordadas mais em detalhe estas implementações bem como todo o trabalho desenvolvido, até ao momento de entrega deste trabalho de Tese, no âmbito do *backoffice* do novo Digitarq. Sendo aqui mencionado um trabalho coletivo da equipa de aplicação, é difícil quantificar o trabalho que diz respeito aos objetivos específicos desta tese. Contudo, o foco foi essencialmente nas permissões dos utilizadores, os formulários de descrição e as intervenções de conservação e restauro.

8.3.1 Dashboards

O *backoffice* da nova versão do Digitarq consiste em duas abas principais: o dashboard do utilizador e o ambiente de trabalho. O dashboard é onde o utilizador pode ver algumas informações relativas ao seu perfil e ao seu mais recente trabalho (documentos recém editados por exemplo). Estes Dashboards mostram informações diferentes para cada perfil do utilizador. Pode ainda, no caso de ser um gestor local ou central (dois perfis de utilizadores existentes), gerir as permissões dos restantes utilizadores. Esta área, até ao momento da entrega desta tese, não foi desenvolvida na sua totalidade pois não afeta o normal funcionamento das funcionalidades essenciais do Digitarq. O que foi desenvolvido foram as áreas de gestor, de modo a se conseguir realizar testes de permissões na plataforma.

8.3.2 Permissões

As permissões dos utilizadores era também um requisito que a DGLAB tinha, com vista a limitar o que um utilizador interno pode fazer ou não. Um utilizador quando é inserido no Digitarq, é lhe dado um ou mais perfis, sendo eles: arquivista, externo, representações digitais, conservação e restauro, revisor, gestor local e central e administrador. A aplicação tem então de estar preparada para limitar certas operações conforme as permissões de cada utilizador. Da parte do *front-end* estas verificações foram feitas usando uma função auxiliar base que trata de toda a lógica necessária, tomando como *input* o tipo de operação a realizar (exemplo: editar) e o tipo de objeto sobre o qual essa operação vai incidir (p.e, um documento, ou ator). Essa função vai verificar ainda o arquivo a que o utilizador pertence, para situações em que seja permitido apenas realizar uma operação sobre objetos do seu próprio arquivo, ou de todos, e ainda, se pode editar todos os objetos, ou só os criados pelo utilizador. Esta foi uma das funcionalidades do *backoffice* realizadas no âmbito desta tese.

The screenshot displays the DIGITARQ workspace interface. At the top, there is a navigation bar with 'Dashboard' and 'Ambiente de trabalho' tabs, a user profile for 'Afonso Esteves', and a 'Terminar Sessão' button. Below this, a toolbar contains 'Ações (4)' and 'Criar' buttons. The main workspace is divided into several sections:

- Top Bar:** 'Ambiente de trabalho' tab, search bar with 'Craveiro Lopes', and 'João Carlos Craveiro Lopes' user information.
- Left Menu (6):** A vertical menu with options: 'Descrição', 'Associações', 'Intervenções', and 'Fechar'.
- Main Card (1):** A large central card titled 'Craveiro Lopes' containing a detailed description of the collection, including levels, identifiers, and titles.
- Hierarchical Description Tree (2):** A table on the right showing 'Descrições Arquivísticas' with columns for 'Código de referência' and 'Título', listing items like 'PT/TT/CRL' and 'PT/TT/PRQ/PMFR11'.
- Digital Representations Tree (3):** A table on the right showing 'Representações Digitais' with columns for 'Nome' and 'Nº ficheiros', listing 'ADP2008PSEC-AG-I(7)' with 0 files.
- Buttons:** 'Imprimir' button at the bottom right of the main card.

Figura 8.13: Ecrã da área de trabalho subdividida pelas diferentes secções

8.3.3 Ambiente de Trabalho

O ambiente de trabalho é onde os utilizadores vão dispende mais tempo e nele podemos encontrar no centro do ecrã uma área de destaque, onde normalmente são abertos os cartões principais. Estes cartões podem estar preenchidos com informações do documento/entidade/representação digital que estamos a ver (8.13) (1) e ainda as árvores hierárquicas do documento (2) ou representações (3). De seguida irão ser elaborados mais estes cartões individualmente. No topo existem dois botões com ações (4) e ainda um sistema de abas da aplicação contendo o objeto (documento, representação ou entidade) que estamos neste momento a manipular (5). Essas abas permitem alterar a informação a ser mostrada e existem para o utilizador conseguir trabalhar com vários objetos em simultâneo, sem o recurso às abas do browser, que resultaria numa experiência de utilização pior, já que obrigaria a sair constantemente da plataforma. Outra vantagem desta prática é que se simula o comportamento da versão anterior do Digitarq a que os utilizadores já estão habituados. Este sistema permite ainda que um utilizador nunca perca o seu ambiente de trabalho já que as informações relativas a que objetos tinha abertos são armazenadas na base de dados relacional. Deste modo, quando se fecha o browser ou se termina sessão, o estado do ecrã permanece igual ao que estava, antes de abandonar a plataforma. À esquerda da área central podemos encontrar ainda um sub-menu (6) com as opções disponíveis, relativas ao objeto que estamos a ver. Por exemplo, se o objeto aberto na tab for um documento, as opções disponíveis são a Descrição do documento, as associações do documento e as intervenções (diagnósticos da conservação e restauro). Por outro lado, se o objeto em destaque for um projeto de representação digital, as opções passam a ser os Metadados do projeto e as Representações digitais.

8.3.4 Descrição do Documento

A descrição do documento consiste num formulário que fica visível num cartão central da área de trabalho quando está um objeto do tipo Documento aberto. Este formulário tem um modo de leitura, que é o modo padrão, e tem ainda um modo de edição, quando um utilizador (com as respetivas permissões), clica no modo editar. Os formulários na sua generalidade têm grande importância no *backoffice* pois estão presentes na grande maioria dos ecrãs. É importante destacar a complexidade desses formulários, que foram projetados para atender às necessidades da DGLAB e substituir os campos de texto livre presentes na versão anterior.

Na mais recente versão, com uma estrutura de dados redesenhada (para uma melhor estruturação dos dados, que permita pesquisas muito mais avançadas dentro da plataforma), a intenção era ter formulários completamente detalhados com todas as opções possíveis de descrição de qualquer objeto. Desde cedo que se preveram as dificuldades de construir estes formulários no *front-end*, então foram feitos templates com base a simplificar o processo. A estrutura base foi feita usando o React Form Hook (mencionado na secção 4.1) e os templates são componentes React que visam abranger todo o tipo de especificidades dos formulários construídos pela equipa de design.

Numa versão inicial, foi pensado as associações (atores, eventos e lugares relacionados) estarem diretamente no formulário do documento mas após alguma análise e teste ao usar a plataforma (ainda em desenvolvimento), foi pensada a solução que está atualmente implementada, que previne que sejam abertos popups para a pesquisa de entidades, que conseqüentemente resultavam em mais popups para a edição dos mesmos, caso se detetasse gralhas ou falta de entidades para relacionar. Desta maneira também se aligeirou o formulário de descrição, que já é consideravelmente longo devido à extensão dos campos a preencher.

8.3.5 Associações

O menu das associações é um menu que é comum aos documentos, entidades e ainda representações digitais. É aqui que são feitas as ligações entre os vários objetos, que se traduzem em ligações entre nós na base de dados de grafos. Estas associações são fundamentais para a compreensão da extensa rede de dados dos arquivos nacionais e permite navegar pela história de uma forma que antes não era possível através do uso de uma base de dados convencional.

Este cartão é constituído por uma parte inicial que permite escolher o tipo de objeto que pretendemos associar e o número de objetos já associados desse tipo, uma secção de filtros a aplicar aos objetos já associados e uma tabela que permite escolher um ou mais desses objetos a desassociar. Por fim, temos ainda um botão que abre outra tabela com todos os objetos que podemos associar. Em ambas as tabelas, em cada linha de um objeto, existe um ícone indicativo de abrir outra tab com os detalhes desse objeto. Essa opção está presente para, no caso de quem está a associar uma entidade, se notar algum erro, ou quiser confirmar se é aquela a entidade certa, poder consultar os detalhes da mesma, numa tab ao lado, podendo sempre voltar ao sítio onde estava. Desta maneira, os requisitos de associar e editar simultaneamente uma entidade ficam cumpridos, evitando uma sequência quase infinita de Popups, com a solução inicialmente pensada.

8.3.6 Intervenções do Documento

As intervenções dizem respeito aos diagnósticos e ações de Conservação e Restauro aplicadas a um documento.

Após terminados os testes de usabilidade, seguiu-se a implementação desta área de Intervenções. Esta secção é composta por duas tabelas, uma para os diagnósticos, e outra para as ações. Entende-se por Diagnóstico, um conjunto de ações realizadas a um documento, num intervalo de tempo. Deste modo, a tabela de cima comporta a lista de Diagnósticos já realizados, enquanto a de baixo, terá a lista de ações de conservação e restauro pertencentes ao diagnóstico selecionado na tabela de cima. Em relação a estes dois tipos de dados, é possível em cada um visualizar os seus detalhes, editar as suas informações e ainda criar novos diagnósticos ou ações.

Foi na realização destes formulários que começaram a surgir grandes dificuldades para atingir o resultado pretendido. Em ambos os formulários encontramos duas abas, uma para os detalhes em si da entidade, e outra para as partes e componentes trabalhadas no diagnóstico em questão. No formulário geral dos detalhes surgiu a necessidade de ter campos com algumas particularidades como: um campo que inicialmente apresenta uma caixa de seleção e que, dependendo da escolha da primeira, pode aparecer ou não, uma segunda caixa, e consecutivamente, uma terceira e quarta. Outro exemplo de campo diferente do que tinha sido desenvolvido até agora, é a junção destas caixas sequenciais com campos múltiplos. Campos múltiplos são essencialmente campos que podem ter várias entradas, como por exemplo, as dimensões. Para esta particularidade de múltiplas entradas usámos uma funcionalidade do Form Hook, chamada de Field Array.

Para além destas dificuldades enfrentadas, a maior delas foi ao desenvolver a aba das Partes e Componentes. De realçar que também existem campos dependentes uns dos outros, e inclusive, campos numa aba que dependem de outros. Desta forma, utilizámos dois *forms*, um em cada aba, e permitimos a sua inter-comunicação através de um componente central que gere ambos. As partes e componentes essencialmente é o local onde se identificam as partes que sofreram algum tipo de restauração no diagnóstico em questão. Aqui teve de ser implementado um sistema de lista de acordeões - um componente usado para abrir/fechar uma determinada secção - de forma hierárquica, isto é, as partes podem ter múltiplos componentes associados, e esses mesmos componentes podem ter também múltiplos sub-componentes. Cada acordeão tem ainda campos associados para especificar alguns detalhes sobre essa parte/componente/subcomponente. A dificuldade aqui foi ter múltiplos campos com a mesma designação, i.e, podemos ter uma parte com os campos: classificação e material, e um componente com os mesmos exatos campos. O Form Hook automaticamente traduz inputs e os seus nomes para um JSON onde mantém o registo de todos os campos de um formulário e os seus respetivos valores. Assim sendo, os nomes dos inputs tiveram de ser prefixados do tipo de objeto e do caminho até chegar ao objeto. Exemplo: o nome de um campo de dimensões que está inserido dentro de um subcomponente “nervo”, seria o seguinte: “parts_estrutura_components_costura_subcomponents_nervo_dimensions”. Desta forma, conseguimos saber ao certo a que campo nos estamos a referir e conseguimos manipular estas strings para montar um JSON correto que podemos mais tarde enviar para a API de dados.

Nesta fase surgiu também outro problema ao utilizar o *FieldArray*. Foram encontrados diversos problemas mas essencialmente era um erro que impedia o bom funcionamento da adição e remoção de campos múltiplos às várias secções de partes e componentes. A abordagem para resolver este problema foi inicialmente reportá-lo à restante equipa em busca de uma resposta. Como não havia sido ainda resolvido, foi requisitada a ajuda de outro membro da empresa, com vários anos de experiência em React e no *hook* em causa. A solução não foi imediata, e foi depois de algumas tentativas que descobrimos a solução (depois de cerca de uma semana). Esta passou por realizar uma versão manual do *FieldArray* manipulando o estado do Form Hook.

8.3.7 Representações Digitais

A área das Representações Digitais é apresentada ao abrir um projeto, representação digital ou ficheiro, através da árvore das representações ou de uma das opções no menu de ações. Esta área assemelha-se ao que temos quando temos um documento aberto, a diferença neste caso é que o cartão central está ocupado com os formulários respetivos (projeto de representação digital, ou uma representação) ou então, no caso de ser um ficheiro, estará presente um visualizador. As opções do sub-menu lateral também são diferentes das opções presentes num documento. No caso de ser um projeto temos a opção de ver o formulário com os metadados deste mesmo projeto e ainda as representações que o compõem. Nesta secção das representações é possível carregar ficheiros. Os ficheiros estão inseridos apenas nas representações digitais, por esse motivo, quando adicionamos ficheiros a partir de um projeto, o que é feito é distribuir os ficheiros (corretamente nomeados) dentro das respetivas representações. Para navegar entre projeto, representação e ficheiro, é utilizada a árvore que se encontra à direita, quando a mesma está aberta.

Quando passamos a visualizar uma representação as opções do menu variam e passam a ser: metadados (semelhante ao que acontece nos projetos), Preservação (implementação planeada para um trabalho futuro), Ficheiros e associações. Na secção de ficheiros é onde podemos ver uma preview dos ficheiros que foram carregadas, bem como a opção de os selecionar para remover. Além disso, podemos ainda dar dois cliques, o que fará com que seja aberta uma nova tab com as informações deste mesmo ficheiro. A aba das associações é muito semelhante ao que acontece nos documentos, com a restrição de apenas ser possível associar documentos. Na vista do ficheiro em si as informações disponíveis são novamente um formulário com informações e um visualizador, idêntico ao que foi implementado no front-office.

Nesta fase foi realizado um grande trabalho de *back-end* para criar as rotas necessárias ao upload e leitura de ficheiros e também do *front-end* para construir três áreas chave ao trabalho das pessoas encarregues pelas representações digitais dos documentos e arquivos nacionais.

8.3.8 Entidades - Ator, Evento, Lugar

As entidades neste projeto são de muito relevo, começando na base de dados de grafos e acabando na interação no *front-end*. Estas estão distribuídas em várias páginas do backoffice, podendo ser acedidas nas suas respetivas tabs, depois de abertas no ambiente de trabalho. É possível editar os

seus detalhes, bem como fazer associações entre entidades e documentos. Para começar, podemos escolher no popup respectivo qual a entidade que queremos abrir. Depois de aberta a tab de uma entidade, deparamos-nos com uma área semelhante àquela que temos quando temos um documento aberto, um cartão central com o formulário da entidade, que podemos ver e editar, e ainda uma subopção de associações. O desenvolvimento destas secções tomou o curso geral, não surgindo grandes obstáculos à sua realização.

No formulário dos lugares foi onde surgiu a única particularidade até ao momento, devido à necessidade de se obter a localização de um lugar através de um mapa. Escrever apenas as coordenadas de um local em caixas de texto não era intuitivo e *user-friendly* suficiente, e por essa razão foi decidido implementar um mapa com o recurso a APIs externas para realizar esta tarefa. Nesta etapa do desenvolvimento foi-me requisitado que fizesse alguma investigação às APIs existentes, as suas características e custos para mais tarde podermos apresentar ao cliente. Ao fazer um levantamento inicial, recolhi quatro delas para uma análise mais detalhada, sendo elas: Google Places API; MapBox API; TomTom API; OpenStreetMap API. A primeira e mais evidente era a da Google. Esta API é de longe a mais completa e com o melhor mapa. Por ser a mais evidente foi também essa que foi inicialmente implementada mas mais tarde, em conversações com o cliente, e com o objetivo em mente de esta plataforma vir a ser uma solução *open-source*, decidiu-se não optar por esta API. Nesta fase também se constatou que seria necessário uma API que fosse responsável por renderizar o mapa, enquanto outra retornasse dados através de uma caixa de pesquisa. A próxima opção na lista era a da MapBox, com a sua solução de API para Javascript, contudo, também esta tinha custos associados. À semelhança da MapBox também a API da TomTom tinha custos, apesar de baixos. Restou uma opção, o OpenStreetMap.

O OpenStreetMap (OSM) é um projeto colaborativo de mapeamento geográfico de código aberto. Ele permite que indivíduos contribuam e editem dados geográficos de qualquer parte do mundo, resultando num mapa detalhado e abrangente. Restava apenas encontrar uma ferramenta que conseguisse mostrar os dados num mapa e a solução encontrada foi a biblioteca Leaflet para React, também ela de código aberto, e relativamente fácil de implementar. Após ter os dados e uma biblioteca que mostre esses dados, faltava ainda encontrar uma API que nos fornecesse resultados da pesquisa de lugares. A mais óbvia é a Nominatim, utilizada no site oficial do OSM. O Nominatim é um serviço de geocodificação e pesquisa de localização baseado em dados do OpenStreetMap. Ele permite converter endereços em coordenadas geográficas (geocodificação) e vice-versa, além de fornecer recursos avançados de pesquisa, como pesquisa por nome de lugar, categorias e tags específicas. O Nominatim é também um projeto de código aberto e pode ser usado gratuitamente, desde que sejam respeitadas as diretrizes de uso e limites estabelecidos pelos fornecedores de serviço. Após encontradas as soluções ótimas e de código aberto, foi realizada novamente a implementação de busca de lugares no formulário respectivo.

8.3.9 Árvores

As árvores estão presentes na secção mais à direita da área de trabalho e consistem na árvore das descrições arquivísticas e das representações digitais. Ambas podem ser mostradas ou escondidas conforme a vontade do utilizador e servem para navegar entre os registos ou as representações. Essencialmente são utilizadas para localizar documentos/representações digitais para serem abertos na secção central, onde se podem ver os detalhes dos mesmos, bem como realizar as outras operações revistas nas secções anteriores.

Ao implementar estas árvores surgiram algumas dificuldades como a paginação e a funcionalidade de guardar o estado da árvore sempre que se muda de tab aberta. Quanto à paginação, foi realizado um grande trabalho, em conjunto com a equipa dos dados, para alterar as rotas da API para suportarem a paginação, uma vez que neste caso não se trata de uma lista linear de coisas, mas sim objetos hierárquicos, i.e, que são compostos por mais objetos, em cadeia. Para atender a estas necessidades e para não sobrecarregar o *front-end*, foi implementada a paginação, com no máximo 10 itens por nível. O que isto significa é que quando se abre um fundo (nível mais acima na hierarquia), e se expande para ver os níveis abaixo, só vão abrir no máximo 10, sendo necessário clicar no botão para ir buscar mais.

Capítulo 9

Conclusão

À medida que esta pesquisa e desenvolvimento alcançam a sua conclusão, é evidente que avançamos significativamente na abordagem dos desafios relacionados à gestão e acessibilidade dos arquivos nacionais em Portugal.

O núcleo deste projeto consistiu na criação de uma interface de pesquisa de arquivos de fácil utilização, com foco na experiência do utilizador. Ao longo deste projeto, foram aplicadas metodologias de design centrado no utilizador e utilizadas tecnologias de ponta para conceber uma plataforma que tem o potencial de transformar a maneira como os utilizadores interagem com o nosso património arquivístico. A entrega bem-sucedida desta sólida solução de *front-end* para a plataforma de pesquisa de arquivos em Portugal representa um marco crucial na melhoria da acessibilidade a esse rico património. Acredita-se que esta plataforma, ao simplificar a navegação e a exploração dos tesouros arquivísticos de Portugal, desempenhará um papel vital na preservação e promoção do legado cultural da nação.

Conforme esta tese chega ao seu termo, olhamos para o futuro com entusiasmo e esperamos que a plataforma concebida contribua para o enriquecimento da compreensão da história e cultura de Portugal. Esta é apenas uma etapa inicial num caminho que visa democratizar o conhecimento e inspirar novas descobertas.

Este projeto representa uma contribuição significativa para o património intelectual e cultural de Portugal. Esperamos que continue a inspirar futuros esforços na área da gestão de arquivos e na promoção do acesso à nossa rica herança histórica.

Bibliografia

- [1] Cristiana Esteves Araújo. *Building the Museum of the Person*. PhD thesis, Universidade do Minho, 2016.
- [2] UK National Archives. UK national archives, 2009.
- [3] U.S. National Archives. U.S. national archives, 2008.
- [4] Ministério da Cultura and Instituto dos Arquivos Nacionais. Isad(g): Norma geral internacional de descrição arquivística: adoptada pelo comité de normas de descrição, estocolmo: Suécia, 19-22 de setembro de 1999, 2002.
- [5] Museu da Pessoa. Museu da pessoa, 2020.
- [6] Museu del Prado. The museo del prado's ontological model, 2016.
- [7] Martin Doerr, Richard Light, and Gerald Hiebel. Implementing the cidoc conceptual reference model in rdf, 2020.
- [8] Europeana. Europeana, 2023.
- [9] Brad Frost. *Atomic Design*. Brad Frost, Pittsburgh, PA, Kansas, 2016.
- [10] M. Gorabbani. react-image-pan-zoom-rotate, 2023.
- [11] ICOM/CIDOC Documentation Standards Group. Definition of the cidoc conceptual reference model, 2003-2011.
- [12] British Museum. British museum, 2020.
- [13] International Council on Archives. *ISAD(G): General International Standard Archival Description*. International Council on Archives, 2000.
- [14] Meta Platforms. Meta platforms, 2021.
- [15] Schema.org. Schema.org, 2015.
- [16] Caixa Mágica Software. Caixa mágica software, 2013.
- [17] Janelle Wong. Atomic design pattern: How to structure your react application, 2017.

Bibliografia

- [1] Cristiana Esteves Araújo. *Building the Museum of the Person*. PhD thesis, Universidade do Minho, 2016.
- [2] UK National Archives. UK national archives, 2009.
- [3] U.S. National Archives. U.S. national archives, 2008.
- [4] Ministério da Cultura and Instituto dos Arquivos Nacionais. Isad(g): Norma geral internacional de descrição arquivística: adoptada pelo comité de normas de descrição, estocolmo: Suécia, 19-22 de setembro de 1999, 2002.
- [5] Museu da Pessoa. Museu da pessoa, 2020.
- [6] Museu del Prado. The museo del prado's ontological model, 2016.
- [7] Martin Doerr, Richard Light, and Gerald Hiebel. Implementing the cidoc conceptual reference model in rdf, 2020.
- [8] Europeana. Europeana, 2023.
- [9] Brad Frost. *Atomic Design*. Brad Frost, Pittsburgh, PA, Kansas, 2016.
- [10] M. Gorabbani. react-image-pan-zoom-rotate, 2023.
- [11] ICOM/CIDOC Documentation Standards Group. Definition of the cidoc conceptual reference model, 2003-2011.
- [12] British Museum. British museum, 2020.
- [13] International Council on Archives. *ISAD(G): General International Standard Archival Description*. International Council on Archives, 2000.
- [14] Meta Platforms. Meta platforms, 2021.
- [15] Schema.org. Schema.org, 2015.
- [16] Caixa Mágica Software. Caixa mágica software, 2013.
- [17] Janelle Wong. Atomic design pattern: How to structure your react application, 2017.