

A EDIÇÃO DIGITAL DE FORAIS MEDIEVAIS PORTUGUESES COM O SUPORTE DE UM SISTEMA DE EDIÇÃO COLABORATIVA EM BASE DE DADOS

THE DIGITAL EDITION OF PORTUGUESE MEDIEVAL MUNICIPAL CHARTERS WITH THE SUPPORT OF A COLLABORATIVE EDITING SYSTEM IN A DATABASE

João Paulo Silvestre *
jpsilvestre@ua.pt

Oswaldo Pacheco **
orp@ua.pt

José Sousa ***
jvcs@ua.pt

Filipa Roldão ****
anaroldao@edu.uletras.pt

Joana Serafim *****
joana.faroserafim@uzh.ch

O artigo aborda a necessidade de um plano de investigação para a edição digital de forais medievais portugueses, destacando três requisitos essenciais. Primeiramente, destaca-se a importância de estabelecer um corpus textual coeso, evidenciando lacunas e desafios interpretativos de maneira consistente. Em segundo lugar, enfatiza-se a necessidade de metodologias e colaboração interdisciplinar, dada a complexidade desses documentos, que exigem contribuições de áreas como História, Diplomática, Paleografia e Linguística. Por fim, ressalta-se a importância de ferramentas digitais para recolha de dados e disseminação de edições digitais em formato aberto. O corpus apresenta desafios, como a interdependência entre documentos e a existência de várias versões dos forais, produzidas ao longo dos séculos. No âmbito do projeto *iForal*, *Forais Medievais Portugueses: uma perspetiva histórica e linguística na era digital*, identificaram-se cerca de 180 forais, com aproximadamente 400 testemunhos sobreviventes. A maioria dos testemunhos encontra-se em latim, representando um desafio linguístico. Desenvolveu-se um *Sistema de Edição Digital Colaborativa* (EDICOLAB) que inclui um módulo de edição, destinado ao trabalho colaborativo dos editores, e um módulo de visualização, destinado ao público geral. Trata-se de uma aplicação em ambiente web, que possibilita que os utilizadores não treinados em marcação XML possam contribuir para a edição

* Centro de Línguas, Literaturas e Culturas, Universidade de Aveiro, Aveiro, Portugal. ORCID: 0000-0002-5980-0075.

** Departamento de Eletrónica, Telecomunicações e Informática - Instituto de Engenharia Eletrónica e Informática de Aveiro, Universidade de Aveiro, Aveiro, Portugal. ORCID: 0000-0002-3098-7163

*** Instituto de Engenharia Eletrónica e Informática de Aveiro, Universidade de Aveiro, Aveiro, Portugal. ORCID: 0000-0001-7163-5116.

**** Centro de História da Universidade de Lisboa, Lisboa, Portugal. ORCID: 0000-0001-8760-6133

***** Universität Zürich, Zúrique, Suíça / Centro de Linguística da Universidade de Lisboa, Lisboa, Portugal. ORCID: 0000-0002-3206-6878.

eletrônica, com intervenções diferenciadas como a correção da transcrição ou a adição de comentários.

Palavras-chave: Forais medievais, Humanidades digitais, Edição crítica.

The article addresses the need for a research plan for the digital edition of Portuguese medieval municipal charters, highlighting three essential requirements. Firstly, it emphasizes the importance of establishing a cohesive textual corpus, consistently highlighting gaps and interpretative challenges. Secondly, it underscores the need for methodologies and interdisciplinary collaboration, given the complexity of these documents that require contributions from areas such as History, Diplomacy, Paleography, and Linguistics. Finally, it highlights the importance of digital tools for data collection and the dissemination of digital editions in an open format. The corpus presents challenges such as the interdependence between documents and the existence of various versions of charters produced over the centuries.

Within the iForal project, *Portuguese Medieval Charters: a historical and linguistic perspective in the digital age*, approximately 180 charters were identified, comprising around 400 surviving testimonies. The majority of these testimonies are in Latin, posing a linguistic challenge. A Collaborative Digital Editing System (EDICOLAB) was developed, including an editing module for collaborative work among editors and a visualization module for the public. It is a web-based application that allows users without XML markup training to contribute to electronic editing with differentiated interventions, such as transcription correction or the addition of comments.

Keywords: Municipal charters, Digital humanities, Critical edition.

•

1. Introdução

Ao estabelecer um plano de investigação sobre forais medievais portugueses identificam-se três necessidades fundamentais. Em primeiro lugar, é essencial delinear um corpus textual coeso e editá-lo de maneira consistente, evidenciando as lacunas na edição e os desafios interpretativos. Em segundo lugar, é necessário estabelecer metodologias e criar as bases para uma colaboração interdisciplinar, uma vez que o estudo desses documentos exige contribuições das áreas de História, Diplomática, Paleografia, Filologia e Linguística e da edição, tanto na crítica textual como nas humanidades digitais. Por fim, a investigação deve apoiar-se em ferramentas digitais que facilitem a recolha de dados, a sua partilha entre os membros da equipa e a disseminação dos resultados de pesquisa num formato de acesso aberto. Embora estes desafios sejam comuns em projetos de edição de manuscritos, há uma particularidade no corpus textual que requer requisitos especiais no sistema de informação: existe uma relação de dependência entre os documentos, ainda não completamente descrita, que a investigação permitirá reconstruir e que deve ser destacada na edição digital.

No âmbito do projeto *iForal, Forais Medievais Portugueses: uma perspetiva histórica e linguística na era digital*, desenvolveu-se um Sistema de Edição Digital Colaborativa (EDICOLAB) que inclui um módulo de edição, destinado ao trabalho colaborativo dos editores, e um módulo de visualização, destinado ao público geral. Trata-

se uma aplicação em ambiente web, que possibilita que os utilizadores não treinados em marcação XML possam contribuir para a edição eletrónica, com intervenções diferenciadas como a correção da transcrição ou a adição de comentários.

O projeto *iForal* encontra-se sediado na Faculdade de Letras da Universidade de Lisboa, tendo como unidades de investigação o Centro de História da Universidade de Lisboa e o Centro de Linguística da Universidade de Lisboa. Trata-se de um projeto interdisciplinar nos domínios da História, da Linguística e das Humanidades Digitais que congrega seis universidades portuguesas, uma universidade no estrangeiro, através de 23 investigadores e seis bolseiros de investigação¹. A Universidade de Aveiro, instituição participante deste consórcio, é a promotora do desenvolvimento do sistema de informação de suporte à edição digital, que inclui também a criação de um glossário de palavras em português.

2. O corpus de forais portugueses medievais

O projeto *iForal* consiste no estudo e edição digital das mais antigas cartas de foral outorgadas pelos primeiros monarcas portugueses às comunidades locais até ao final do reinado de D. Afonso III, isto é, até 1279. Nestas cartas, definiam-se normas de convivência entre habitantes e entre estes e os poderes locais e os reis em âmbitos tão diversos como a justiça, o pagamento de tributos ou impostos, a hierarquia e a convivência social, a defesa militar, a transação de bens e animais, a delimitação do espaço concelhio, entre outros (Costa, 1971).

Os forais produzidos e copiados neste período encontram-se maioritariamente escritos em língua latina, tendo conhecido, ao longo dos séculos seguintes, novas versões não só em latim, mas também em língua vernacular. Por este motivo, o projeto abarca uma cronologia de estudo mais lata, que se estende até ao final do século XV, de modo a compilar os testemunhos sobreviventes destes forais, antes mesmo da renovação administrativa e burocrática a que a chancelaria régia estará sujeita a partir do reinado de D. Manuel (Roldão & Serafim, 2021).

Estes testemunhos foram produzidos, ao longo dos séculos, por circunstâncias diversas, por iniciativa quer de monarcas, quer das comunidades e poderes locais a que se destinavam. Por um lado, a transição de reinado implicava, na maior parte dos casos, a confirmação por escrito das concessões e privilégios outorgados pelo rei ou reis anteriores, conferindo maior legitimidade ao exercício do poder pelo novo rei. As cartas de foral são disso um bom exemplo. Por outro lado, no seio das comunidades a que se destinavam os forais, revelava-se, periodicamente, a necessidade de renovar o testemunho vigente do foral da terra, por motivos relacionados com a perda ou destruição da carta de foral, pela deterioração das existentes ou mesmo porque eram já incompreensíveis os seus conteúdos, em latim e em letra manuscrita difícil de ler. Alegando tais motivos, pedia-se ao rei uma nova cópia do foral ou, em reunião da câmara do concelho, decidia-se atribuir a um escrivão, nomeadamente a tabeliães públicos, a produção de novos testemunhos autênticos do foral, agora já não em latim, mas em vernáculo.

¹ Mais informações sobre o projeto podem encontrar-se em <https://iforal.hypotheses.org/>.

Por este motivo, cada foral terá conhecido várias versões, produzidas em instâncias de escrita e sob critérios de execução diferenciados e em registos de língua com variações ao longo dos séculos. Todos estes testemunhos se revelam de enorme interesse histórico e linguístico para o projeto, não só porque permitem definir as circunstâncias de transmissão ao longo de quase quatro séculos de uma mesma carta de foral, assim como permitem estabelecer comparações com os forais de outras terras.

Esta comparação torna-se ainda mais aliciante quando procuramos conhecer as lógicas de atribuição de cartas de foral às comunidades. Na verdade, existiram três modelos principais e orientadores para a redação de uma carta de foral pelo rei, adotados no século XII para muitas localidades portuguesas: dois modelos ibéricos e um português. O mais antigo é o modelo de foral da cidade leonesa de Salamanca, introduzido em Portugal pela primeira vez em Numão, em 1130, e que se espalhou principalmente pelas regiões norte e noroeste do território. Posteriormente, surgiu o modelo castelhano de Ávila, perdido e apenas conhecido através dos seus “descendentes”, nomeadamente o foral de Évora, atribuído a esta cidade em 1166. Este modelo disseminou-se principalmente nas regiões do centro-sul e sul de Portugal. Por último, temos o tríplice foral de Coimbra-Santarém-Lisboa, de origem portuguesa, adotado por essas cidades em 1179 e presente principalmente na região centro e litoral do território (Reis, 2002, 2006).

Outros forais agrupam-se em famílias de menor expressão ou, eventualmente, não estão integradas em nenhuma família, sendo ainda necessário confirmar as informações de algumas famílias de forais, integrar testemunhos e ainda definir a situação de um conjunto significativo de forais. De qualquer modo, sabe-se, pela própria descrição da carta de foral e pela colação dos testemunhos, a forma como se deu a transmissão destes textos. O texto atribuído à primeira localidade, ou seja, o foral-modelo era transmitido a outras localidades, dependendo do seu estatuto e situação geográfica, entre outros fatores. No entanto, este texto-modelo não era copiado integralmente: apesar de as principais cláusulas constarem na grande maioria dos forais, havia, com bastante frequência, uma adaptação do conteúdo ou até da redação às características da localidade que estava a receber o foral (aspetos geográficos, produtos de comércio, privilégios militares, etc.). Por sua vez, estes forais que recebem o foral-modelo — forais-variante — podiam ser atribuídos a outras localidades, tornam-se, ele próprios, também em forais-modelo. Vejamos um exemplo: o foral de Salamanca foi atribuído a Numão (1130), Freixo de Espada à Cinta (1152), Mós (1162), Trancoso (1157–1169), Linhares (1169) e Guarda (1199); por sua vez, apenas os forais de Numão e Mós não foram atribuídos a outras comunidades; por exemplo, o foral de Trancoso foi o modelo concedido a dez outras localidades, entre as quais Celorico (1157–1169), que por sua vez serviu de modelo a Vila Boa do Mondelo em 1216 (Reis, 2002, 2006). Temos, portanto, uma sequência de forais-modelo que se convertem em forais-variante e esta interdependência é, pois, bem visível com o cotejo dos textos.

O projeto *iForal* identificou cerca de 180 forais, num total de aproximadamente 400 testemunhos de cartas de foral sobreviventes no período em apreço, considerando, como acima descrito, os mais antigos forais régios atribuídos até ao final do reinado de D. Afonso III e as suas cópias latinas e vernaculares até ao final do século XV.

Este corpus apresenta características que o tornam um desafio para o seu estudo. Por um lado, apenas 5% dos testemunhos coligidos correspondem a cartas de foral originais, já que se perderam os documentos autênticos emanados pelas chancelarias dos reis outorgantes destas cartas e que, por consequência, teriam sido enviados às comunidades. Na maioria dos casos, as cartas de foral sobreviveram no corpo do texto de traslados e de confirmações feitos por reis posteriores, ou mesmo no interior de uma carta de foral que diz copiar o foral de outra terra, hoje perdido. Por estes motivos, a maioria dos testemunhos encontra-se copiada em livro, designadamente, nos livros da chancelaria régia dos primeiros monarcas.

Por outro lado, dos 400 testemunhos coligidos, apenas cerca de 10% correspondem a documentos em vernáculo, prevalecendo, por isso, os forais em língua latina, uma vez que esta era ainda a língua utilizada na chancelaria para a redação dos documentos oficiais, apesar de o português circular como língua de comunicação oral já há muito tempo. Destes forais em vernáculo, cinco (num total de 13 testemunhos) chegaram até nós exclusivamente sem um texto correspondente em latim². De qualquer modo, podemos afirmar que 90% do corpus se encontra disponível apenas para conhecedores da língua latina.

Finalmente, a relação que se assume existir entre testemunhos, no seio de uma transmissão textual dentro de uma família de forais e/ou sob um mesmo modelo de foral, está por confirmar e reinterpretar quer no âmbito da historiografia portuguesa, quer nas áreas da crítica textual e da filologia latina e portuguesa.

Na verdade, estas características não têm condicionado apenas o estudo destes manuscritos, mas também as iniciativas de edição que ao longo dos séculos têm vindo a incidir sobre os mais antigos forais régios.

Com efeito, foi Alexandre Herculano, no século XIX, que pioneiramente apresentou, nos volumes dos *Portugaliae Monumenta Historica*, (PMH) dedicados às leis e costumes, a edição crítica da maioria dos forais que compõem o nosso corpus de outros outorgados por outros senhores que não o rei (Herculano, 1856–1977). Nessa edição, Herculano apresentava os testemunhos sobreviventes, algumas notas de crítica interna e externa sobre os mesmos, um aparato de variantes, assim como, nos casos possíveis, a versão em vernáculo dos testemunhos latinos. Contudo, o esforço de compilação de Herculano não teve continuidade nas edições que os forais vieram a conhecer posteriormente, já no século XX. Talvez também por desinteresse historiográfico sobre o próprio documento/monumento “foral” e os seus significados, os forais foram tendencialmente editados no seio de publicações mais vastas que incidiram na edição completa dos livros de chancelaria em que se encontravam copiados. Ou, do mesmo modo, editados apenas no âmbito de publicações que privilegiaram a edição dos documentos emanados por um rei. Perdia-se, assim, a lógica de edição de forais, enquanto corpo documental autónomo, composto por múltiplos testemunhos, conservados em fundos e em arquivos diferenciados, para se atender à lógica da unidade de instalação ou do próprio reinado a que o foral estaria associado.

² Trata-se dos forais de Almada, Condado, Guardão, Póvoa d’El-Rei e Telões de Aguiar.

Paralelamente, a profusão de edições de forais por organismos de poder ligados às comunidades, por eruditos locais ou até por investigadores em trabalhos científicos, ainda que assumam importância num panorama global, tendem a acentuar a dispersão editorial dos forais medievais (Coelho, 2011).

Mais recentemente, uma nova edição dos PMH, em dois volumes, da autoria de António Matos Reis veio possibilitar voltarmos a dispor de um corpus de forais medievais como objeto autónomo e relacional, ainda que não se trate de uma edição crítica (Reis, 2019, Reis, 2023).

Em qualquer dos casos acima mencionados, as edições disponíveis não esgotam o universo de testemunhos que é possível identificar para o período medieval e, mais importante, encontram-se em papel e foram preparadas para atender a esse formato. Na verdade, não existia, até agora, uma edição digital de forais medievais, ainda que seja possível consultar algumas das publicações acima mencionadas na sua versão digitalizada.

Ora, esta circunstância, aliada à complexidade das relações de dependência que se conhecem e se intuem nestes documentos, assim como à necessidade de identificar e analisar criticamente cada testemunho sobrevivente, evidencia, com clareza, a importância de criar um sistema de informação capaz de atender a estes desafios. Na verdade, a resposta que o projeto *iForal* propõe para este corpus é, antes de mais, uma resposta adaptável a outros conjuntos documentais que necessitem de uma edição digital.

3. A edição digital

Uma edição académica digital de um conjunto de textos interrelacionados requer informações que descrevam a estrutura complexa dos textos, o seu conteúdo e a forma como foram transmitidos no formato manuscrito, resultando numa apresentação clara destes elementos ao leitor.

O desafio, no que respeita à organização do trabalho editorial, é permitir que os membros da equipa responsáveis por criar cada nível de informação possam registar esses dados de maneira autónoma, contribuindo assim diretamente para a edição digital. A automatização desse processo é crucial, evitando a dependência de um membro da equipa dedicado exclusivamente à codificação de dados previamente elaborados.

Existem poucas ferramentas que dispensam o conhecimento explícito de linguagem de marcação, o que acarreta desafios para a participação efetiva dos membros da equipa na codificação de uma edição digital em XML TEI, particularmente quando essa edição envolve um aparato crítico.

Para uma edição digital não é suficiente a disponibilização do fac-símile do manuscrito ou de uma transcrição do texto. Distingue-se também das edições impressas que foram convertidas para o formato digital através de digitação ou outro processo automático, conhecidas como edições digitalizadas.

As edições concebidas como digitais devem oferecer soluções que não se encontram nas edições tradicionais impressas, como meios de leitura, exploração e análise do texto. Adicionalmente, podem incluir identificadores das relações com outros textos.

As informações de carácter filológico a incluir na edição digital dão conta de dificuldades relacionadas com a escrita, leitura e interpretação. No caso dos forais medievais, a edição depara-se com problemas de deterioração do suporte e legibilidade, variação nas convenções de representação escrita, bem como problemas de omissões e erros introduzidos pelos escrivães no processo de cópia. Os editores devem avaliar cuidadosamente esses erros e tomar decisões fundamentadas sobre se os corrigem ou mantêm na versão editada, procurando equilibrar a preservação da autenticidade do documento com a sua compreensibilidade para os leitores contemporâneos. Na edição digital, os editores tornam visíveis as suas intervenções no sentido de modernizar a linguagem, bem como a interpretação de passagens obscuras ou ilegíveis.

Os forais frequentemente fazem referência a outros documentos, indivíduos ou lugares. Os editores estabelecem a relação entre esses elementos e anotam o documento para fornecer contexto histórico, explicações e esclarecimentos, sem alterar o texto original. A relação com outros forais contribui para criar uma representação mais completa do documento que está a ser editado. Como de alguns forais existem várias cópias, cada uma com variações inerentes ao processo de transmissão, os editores podem necessitar de comparar as diferentes versões para justificar uma solução de edição.

Por último, a qualidade filológica da edição digital não reside apenas na transcrição, mas também na adição de informação codificada com metadados que possibilitam a utilização em investigações futuras, com um alto grau de compatibilidade com os sistemas informáticos, contribuindo assim para a preservação digital e acessibilidade.

4. Exemplos de projetos de edição digital

Uma análise de sistemas de informação de suporte à edição, resultantes de projetos de investigação, permite constatar a preferência por tecnologias de base de dados que priorizem a persistência e manipulação de dados no formato XML. Por tratarmos da edição de manuscritos, referimos os projetos *CEED*, *Patrimonivm Editor* e *EVT* (Sousa, 2023, pp. 10–12).

O *Cooperative Web-Based Editor for Critical Editions (CEED)* é uma solução desenvolvida no contexto do projeto *PhiBor*, focado no estudo de manuscritos medievais de filosofia árabe (Marotta, 2021). O objetivo principal deste sistema de informação é otimizar o processo de edição crítica e facilitar a colaboração entre utilizadores durante esse processo. A edição colaborativa concorrente é alcançada por meio do desenvolvimento de um mecanismo que permite a divisão do documento em partes atribuídas a diferentes utilizadores. Essa abordagem inova ao combinar eficiência com a simplicidade do *locking*, evitando a perda de dados numa edição concorrente.

Do ponto de vista tecnológico, o sistema é construído utilizando o *framework* Angular e a linguagem Typescript. A comunicação com o *backend* segue a arquitetura Representational State Transfer (REST), utilizando objetos *JavaScript Object Notation* JSON. No *backend*, a base de dados empregada é o BaseX, otimizado para a persistência de XML, e a linguagem XQuery é utilizada para a transformação e consulta de dados estruturados e não estruturados. O armazenamento de imagens segue os padrões do *International Image Interoperability Framework*, assegurando a interoperabilidade entre

sistemas e possibilitando a visualização ao lado do editor de texto (Sousa, 2023, pp. 10–12).

O *Patrimonivm Editor* é um sistema de informação desenvolvido no âmbito do projeto *Patrimonivm*, com o intuito de realizar um estudo multidisciplinar sobre o papel político, social e económico das propriedades dos imperadores romanos (Broux, 2019). Para além de simplificar o processo de edição colaborativa entre os investigadores, este sistema apresenta uma nova funcionalidade em comparação com projetos anteriores.

Para além da edição de texto, suporta a georreferenciação dos locais mencionados nos manuscritos. Apesar de os autores terem um foco na adoção do TEI, há uma preocupação em oferecer suporte a marcações não sistematizadas. Para tal, disponibilizam um conversor de marcação que reconhece automaticamente a simbologia não sistematizada e a converte posteriormente para TEI. O *frontend* do sistema é renderizado do lado do servidor utilizando XQuery e XSLT, devolvendo ao cliente a página pronta para ser apresentada em HTML, Javascript e CSS (Sousa, 2023, pp. 12–13).

O EVT é uma aplicação web para visualização de edições críticas em TEI, desenvolvida no âmbito do projeto *The Digital Vercelli Book*, que tem como objetivo a edição digital do livro de Vercelli e a disponibilização da edição resultante (Rosselli del Turco, 2019; Rosselli del Turco et al., 2014). O propósito desta aplicação é proporcionar aos editores a capacidade de gerar automaticamente uma página web, onde seja possível visualizar o documento editado através de um visualizador TEI. Esta abordagem inovadora visa simplificar a maneira como as edições são apresentadas aos leitores, uma vez que em outros contextos é comum haver um desenvolvedor web cuja única função é converter manualmente um documento TEI para HTML.

A aplicação é desenvolvida utilizando o framework *Angular* e utiliza transformações XSLT para criar um arquivo HTML, pronto para ser renderizado no navegador, a partir de um arquivo TEI XML. De forma geral, as transformações XSLT possibilitam a conversão parcialmente automática de uma linguagem baseada em XML para outra. A interface do sistema procura simular um livro que pode ser folheado. Na data em que a informação pública sobre este projeto foi consultada, os autores desenvolvem uma nova versão que permitirá aos utilizadores personalizar a interface, carregar um aparato crítico e criar marcadores em cada página (Sousa, 2023, pp. 14–15). O projeto CEED é o que implementa o maior número de funcionalidades necessárias para a edição crítica de forais medievais. No entanto, a falta do glossário e de um visualizador do documento torna essa solução incompleta, indicando a necessidade de desenvolver um novo sistema de informação que integre editor, visualizador e glossário.

Nenhum dos projetos estudados implementa um modo de edição colaborativa totalmente concorrente, optando por mecanismos de bloqueio para garantir a consistência dos dados. Esta escolha é justificada pelo nível de complexidade envolvido na edição concorrente, que só se torna justificável em projetos que exigem a edição simultânea por várias pessoas.

Um problema identificado em todos os projetos é a falta de documentação relacionada à implementação dos sistemas de informação. Preferencialmente, descrevem-se as características do texto e os critérios de edição, deixando de fora aspetos técnicos

relevantes para a engenharia de software. Uma exceção é o projeto EVT, que disponibiliza o código-fonte e documentação detalhada sobre a implementação.

5. A codificação da edição digital dos forais

A prática predominante nas Humanidades Digitais envolve a utilização da linguagem XML-TEI, desenvolvida pela Iniciativa de Codificação de Texto (TEI, 2007), devido às vantagens que proporciona na criação de edições com múltiplos níveis de informação. Em primeiro lugar, a linguagem XML é uma forma de marcação que descreve o texto, a sua estrutura lógica e os aspetos semânticos, formalizando o conhecimento do editor sobre o texto a ser anotado e facilitando a sua interpretação pelos leitores. Esta difere de marcações mais simples, como o HTML, que apenas definem como o texto deve ser processado.

A adoção desta linguagem estabelece uma base comum de trabalho para os colaboradores de um projeto e, do ponto de vista dos sistemas de informação, maximiza a interoperabilidade entre sistemas. Uma vantagem substancial é que um grupo de trabalho não fica vinculado a uma solução de software específica, permitindo a migração para uma nova ferramenta sem a necessidade de reeditar os documentos.

Dado que os forais representam apenas um entre vários tipos de textos manuscritos produzidos durante o período medieval, torna-se essencial desenvolver um sistema de edição e visualização com base na linguagem TEI, que ofereça versatilidade e adaptabilidade. Este sistema deve ter a capacidade de abranger diversos tipos análogos de textos, incorporando funcionalidades que permitam aos utilizadores personalizar o sistema conforme as exigências tanto da edição quanto do próprio texto. Estas funcionalidades podem incluir a definição de categorias de texto, variações de anotações e preferências de apresentação.

A marcação TEI revela-se suficientemente flexível para acomodar as características da maioria dos manuscritos medievais, possibilitando a descrição de aspetos tão diversos como fenómenos linguísticos complexos, informações topográficas ou a materialidade dos forais enquanto objetos arquivísticos.

No entanto, a emergência de novos projetos de edição suscita a interrogação sobre se o padrão TEI deve ser aplicado em textos com requisitos particularmente específicos ou se é necessário desenvolver outros padrões de marcação. Um exemplo notório desse debate é a iniciativa do *Charters Encoding Initiative* (CEI), que procura estabelecer um padrão específico para textos legais medievais. Este grupo sustenta que as soluções oferecidas pela TEI são inadequadas ou excessivamente complexas para descrever aspetos específicos desses manuscritos. Uma das situações mais complexas é a marcação e descrição dos elementos de validação e certificação dos manuscritos, como assinaturas e selos de diversos tipos, assim como a autorização e certificação dos documentos. Este grupo de trabalho iniciou as suas atividades em 2004, analisando a versão P4 da linguagem TEI. Embora a versão atual, P5, já permita abordar alguns aspetos relacionados com a descrição dos elementos de autenticação, os membros do grupo consideram que, de certa forma, ainda é insuficiente para as suas necessidades específicas.

O modelo de codificação dos forais foi desenvolvido utilizando uma versão personalizada das Diretrizes P5 da Iniciativa de Codificação de Texto (TEI). O cabeçalho do TEI XML representa a primeira secção do documento, contendo informações sobre o texto e o seu formato. Este inclui metadados como título, autor, data de criação e idioma, sendo essenciais para que programas de processamento de texto possam reconhecer as características do documento e interpretá-lo corretamente. O cabeçalho, conforme ilustrado na Figura 1, é constituído por diversos componentes, tais como <fileDesc>, <encodingDesc>, <profileDesc>, <xDATA> e <revisionDesc>, cada um desempenhando um papel específico na descrição do documento.

Figura 1. Exemplo de marcação do cabeçalho de um documento TEI-XML.

```
<teiHeader>
  <fileDesc>
    <titleStmnt>
      <title>Foral de Guarda</title>
      <editor role="transcriber">Rui Pedro Neves</editor>
      <editor role="transcriber">Catarina Coelho</editor>
      <funder>FCT</funder>
    </titleStmnt>
    <publicationStmnt>
      <publisher>Centro de História da Universidade de Lisboa</publisher>
      <pubPlace>Lisboa</pubPlace>
      <date>Jan 12 1411</date>
      <availability>
        <licence target="http://creativecommons.org/licenses/by-sa/3.0/">
          Creative Commons Attribution-ShareAlike (CC BY-SA)
        </licence>
      </availability>
    </publicationStmnt>
    ...
  </fileDesc>
</teiHeader>
```

No interior do componente <msDesc>, representado na figura 2 e relevante para a edição de forais, são empregues outros componentes para codificar informações relacionadas à descrição do manuscrito, abordando aspetos como a localização espacial, características físicas e contexto histórico ou cultural. Dentro destes, apenas o componente <msIdentifier> é considerado obrigatório. Este componente desempenha um papel central ao identificar de forma única o manuscrito, fornecendo dados essenciais como números de catálogo, cota, ou outros identificadores que garantam a precisão e singularidade na referência ao documento em questão.

Figura 2. Exemplo da informação no componente ``<msDesc>``

```

<teiHeader>
  <fileDesc>
    <titleStmnt>
      <title>Foral de Guarda</title>
      <editor role="transcriber">Rui Pedro Neves</editor>
      <editor role="transcriber">Catarina Coelho</editor>
      <funder>FCT</funder>
    </titleStmnt>
    <publicationStmnt>
      <publisher>Centro de História da Universidade de Lisboa</publisher>
      <pubPlace>Lisboa</pubPlace>
      <date>Jan 12 1411</date>
      <availability>
        <licence target="http://creativecommons.org/licenses/by-sa/3.0/">
          Creative Commons Attribution-ShareAlike (CC BY-SA)
        </licence>
      </availability>
    </publicationStmnt>
    ...
  </fileDesc>
</teiHeader>

```

6. Objetivos e requisitos

O sistema de informação desenvolvido, designado *Sistema de Edição Digital Colaborativa* (EDICOLAB), visa simplificar a colaboração no processo de construção de uma edição crítica multidisciplinar, permitindo também a publicação das edições críticas numa página web acessível ao público. O sistema procurou cumprir os seguintes objetivos:

- Implementar um editor de texto TEI capaz de validar o código-fonte a nível sintático, facilitando o processo de edição tanto para utilizadores com diferentes níveis de experiência.
- Possibilitar a edição colaborativa de um documento, permitindo que vários utilizadores contribuam para a construção da edição de forma simultânea.
- Gerir as permissões de leitura e escrita dos diferentes utilizadores no modo edição, garantindo um controlo eficaz sobre quem pode aceder e modificar o documento.
- Implementar um visualizador TEI que construa uma representação visual a partir das etiquetas XML.
- Implementar um visualizador de imagens de alta resolução do documento, permitindo que os utilizadores se foquem em detalhes específicos.
- Possibilitar a construção de um glossário colaborativo que possa ser consultado durante a edição ou leitura de um documento, facilitando a compreensão de termos específicos.
- Desenhar uma arquitetura baseada em tecnologias web que assegure a escalabilidade horizontal do sistema, permitindo que este cresça em resposta à crescente utilização.

A edição colaborativa é uma funcionalidade associada a uma aplicação de edição de texto, que permite que duas pessoas ou mais modifiquem o mesmo documento. É utilizado um mecanismo de *locking*, que faz com que quando um utilizador abre um documento no modo de edição este fique com a exclusividade de o editar. Só quando o utilizador fecha o documento é que outro pode tomar o seu lugar e fazer alterações.

O visualizador compreende a leitura do documento e a visualização das imagens fac-símile. O visualizador do documento é uma aplicação que permite aos utilizadores visualizar um documento para além do seu código-fonte. Normalmente, os visualizadores interpretam as etiquetas XML e dão-lhes um aspeto gráfico. O visualizador das imagens fac-símile é um visualizador de fotografias de alta resolução do documento original.

O glossário, neste sistema, é entendido como uma lista de palavras que explicam o significado de termos através de definições, equivalências ou contextos textuais. Normalmente, numa edição crítica construída recorrendo a software de edição de texto, esta lista aparece no fim da edição de forma não sistematizada. Contudo, uma vez que o TEI prevê etiquetas próprias para a codificação de glossários, é possível sistematizar esta informação, enriquecendo assim a experiência de leitura de uma edição.

Da análise dos objetivos resultou uma lista de requisitos funcionais e não funcionais com diferentes prioridades. Os requisitos funcionais referem-se a funcionalidades do sistema, enquanto os não funcionais evidenciam comportamentos e características do sistema.

Entre os requisitos mais relevantes, destaca-se a introdução de um editor de texto WYSIWYG permite a marcação direta do corpo do documento. Esta é uma abordagem prática para utilizadores mais experientes que desejam manipular o código-fonte TEI de forma direta.

Figura 3. Editor de texto (Foral de Trancoso, 1217)

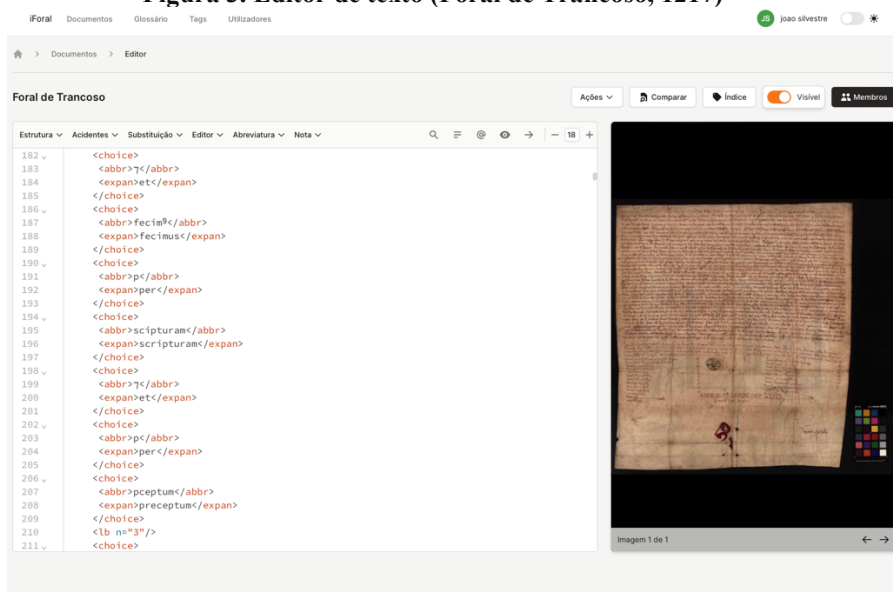
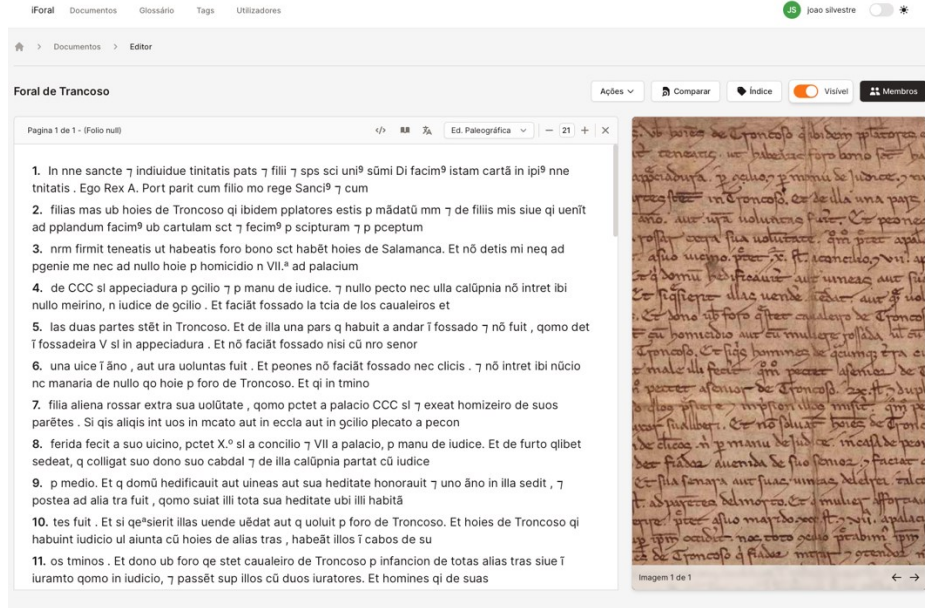


Figura 4. Visualizador de texto, edição paleográfica (Foral de Trancoso, 1217)

Ainda assim, a exigência de suporte para edição direta e indireta do código-fonte é assegurada pela presença de menus, que facilitam a personalização de etiquetas TEI para utilizadores menos experientes. Foi também considerado prioritário a criação de um formulário para a recolha de metadados do cabeçalho de uma edição TEI.

Figura 5. Formulário do cabeçalho TEI (Foral de Trancoso, 1217)

Editar cabeçalho
 Título da Obra: Foral de Trancoso
 Editores: Rui Pedro Neves (transcriptor), Catarina Coelho (encoder), Vasco Sousa (Selegonar papel)
 Financiadores: FCT

Publicação
 Produtor: Centro de História da Universidade de Lisboa
 Local de Publicação: Lisboa
 Data de Publicação: 10/12/2022
 Licença: Creative Commons Attribution-ShareAlike (CC BY-SA)

Descrição da Fonte
 País: Portugal
 Instituição: Arquivo Nacional Torre do Tombo
 Local: Lisboa
 Repositório: Feltes da Coroa
 Idno: PT/TT/F/CI/001427
 Data de Origem: 01/10/1217 - 23/01/2024

A implementação de um mecanismo de pesquisa no texto existente na base de dados do sistema é igualmente um dos requisitos basilares deste projeto. Esta funcionalidade dá aos utilizadores a capacidade de filtrar os documentos por intervalo de datas, título do foral e língua em que está redigido. Para além da pesquisa básica, existe também um modo avançado de pesquisa de texto, que possibilita a inserção de palavras em pesquisa simples ou em contexto.

A compilação de um glossário TEI através de formulário garante a interoperabilidade do glossário construído no sistema de informação entre diferentes plataformas, promovendo a consistência e a eficaz utilização dos dados.

Assegurar a autenticação dos utilizadores é importante para possibilitar a colaboração em documentos, com base em permissões individuais. Isso permite que os administradores possam convidar novos membros para edição e impede que os membros da equipa alterem documentos dos quais não são editores.

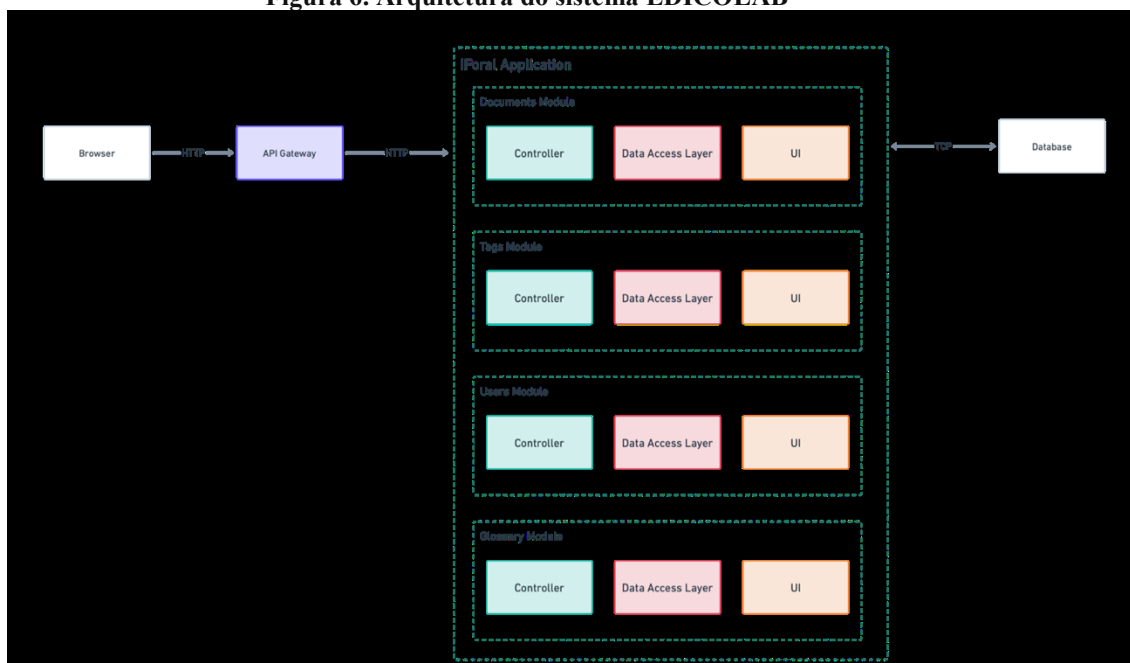
Por último, a implementação de um visualizador permite a consulta de imagens em alta resolução durante todo o processo de edição, simplificando a tarefa dos utilizadores e eliminando a necessidade de alternância entre duas aplicações.

No que diz respeito aos requisitos não funcionais, que não são o foco do presente trabalho, foram considerados requisitos que asseguram a usabilidade do sistema, a segurança e o desempenho.

A solução de arquitetura foi desenhada para dar resposta às necessidades do sistema de edição colaborativa, e também para que fosse possível implementá-la dentro do prazo estipulado para o projeto iForal. Trata-se de uma arquitetura monolítica modular, que agrega todas as funcionalidades num único executável que pode ser carregado para um servidor, eliminando a necessidade de comunicação na rede entre as diversas funcionalidades do sistema. Esta arquitetura é indicada para projetos com uma equipa de desenvolvimento pequena e com prazos de conclusão curtos.

Os detalhes técnicos do diagrama apresentado na figura 6 podem ser consultados em (Sousa, 2023, pp. 41–43).

Figura 6. Arquitetura do sistema EDICOLAB



7. Conclusão

Apesar de existirem várias soluções para a edição digital de manuscritos, justifica-se o desenvolvimento de um sistema de informação que seja adaptado às características documentais e linguísticas específicas dos forais medievais portugueses. Tal sistema deve também considerar a composição das equipas de investigadores dedicadas ao estudo e à edição desses documentos.

O contributo das tecnologias da informação incide sobre três áreas: a colaboração da equipa multidisciplinar, a codificação de textos e a experiência do utilizador.

Sobre a colaboração da equipa multidisciplinar, o sistema de informação deve facilitar a coordenação e distribuição de tarefas entre os colaboradores, promovendo uma comunicação eficiente. Além disso, a capacidade de introduzir informação de forma autónoma assegura que cada membro da equipa possa contribuir de acordo com a sua especialidade.

A codificação do texto de manuscritos medievais é complexa, pois o conjunto documental apresenta testemunhos de diferentes etapas da evolução da língua e das convenções de representação escrita. Embora existam normas de codificação, como a TEI, a aplicação é complexa e exige treino especializado. Neste contexto, é essencial desenvolver soluções tecnológicas que simplifiquem o processo de codificação e transcrição, tornando-o mais intuitivo e acessível aos editores e investigadores.

Financiamento: Esta pesquisa foi elaborada no âmbito do projeto *iForal, Forais Medievais Portugueses: uma perspetiva histórica e linguística na era digital* (PTDC/HAR-HIS/5065/2020), financiado por fundos nacionais pela Fundação para a Ciência e Tecnologia (Ministério da Ciência, Tecnologia e do Ensino Superior).

Referências

- Broux, Y. (2019). PATRIMONIVM: Geography and economy of the imperial properties in the Roman world. <https://patrimonium.huma-num.fr>
- Coelho, M. H. da C. (2011). Municipal power. In J. Mattoso (Dir.), M. de L. Rosa, B. Vasconcelos e Sousa, & M. J. Branco (Eds.), *The historiography of medieval Portugal c. 1950-2010* (pp. 209–230). Instituto de Estudos Medievais.
- Costa, M. J. de A. (1971). Forais. In J. Ferrão (Dir.), *Dicionário de história de Portugal* (vol. 2, pp. 279–281). Iniciativas Editoriais.
- Herculano, A. (Ed.). (1856–1977). *Portugaliae monumenta historica a saeculo octavo post Christum usque ad quintum decimum, iussu Academiae Scientiarum Olisiponensis: Leges et consuetudines*. Biblioteca Nacional Digital. <https://purl.pt/12270> ei
- Marotta, D. (2021). Tracing a digital critical edition. The experience of the PhiBor project. *Umanistica Digitale*, 5(10), 89–114. <https://doi.org/10.6092/issn.2532-8816/12623>
- Reis, A. M. (2002). *Origens dos municípios portugueses* (2.^a ed.). Livros Horizonte.
- Reis, A. M. (2006). *História dos municípios (1050-1383)*. Livros Horizonte.
- Reis, A. M. (Ed.). (2019). *Leges et consuetudines: Forais e cartas de povoamento (1^a parte, 1050-1248)*. Academia das Ciências de Lisboa.
- Reis, A. M. (Ed.). (2023). *Leges et consuetudines: Forais e cartas de povoamento (2^a parte, 1244-1385)*. Academia das Ciências de Lisboa.

- Roldão, F., & Serafim, J. (2021). Os mais antigos forais régios portugueses: Uma proposta de estudo e de edição. In R. Martínez Peñín, & G. Caveró Domínguez (Coords.), *Poder y poderes en la Edad Media* (pp. 375–386). Sociedad Española de Estudios Medievales. https://medievalistas.es/wp-content/uploads/2021/08/poder-y-poderes_SEEM-16-WEB.pdf
- Sousa, J. (2023). *Sistema integrado de suporte à edição eletrónica de forais medievais portugueses e de desenvolvimento de um glossário* [Tese de mestrado, Universidade de Aveiro].
- TEI – Text Encoding Initiative. (2007). <https://tei-c.org/>
- Rosselli del Turco, R. (2019). Designing an advanced software tool for Digital Scholarly Editions: The inception and development of EVT (Edition Visualization Technology). *Textual Cultures: Texts, Contexts, Interpretation*, 12, 111–191. <https://doi.org/10.14434/textual.v12i2.27690>
- Rosselli del Turco, R., Pietro, C. D., Julia, K., & Raffaele, M. (2014). EVT - Edition Visualization Technology.

[recebido em 22 de janeiro de 2024 e aceite para publicação em 01 de julho de 2024]