

Do manuscrito à edição digital

Os percursos de uma investigação interdisciplinar
sobre forais medievais portugueses

From Manuscript to Digital Edition

Tracing Interdisciplinary Research on Portuguese Municipal Charters

FILIPA ROLDÃO*

JOÃO PAULO SILVESTRE**

JOANA SERAFIM***

RESUMO O presente artigo visa apresentar as fases de conceção, experimentação e implementação de um sistema de edição digital direccionado a um *corpus* de manuscritos não literários, os forais régios portugueses. O *corpus* deste trabalho contempla as cartas de foral outorgadas

* <https://orcid.org/0000-0001-8760-6133>

Universidade de Lisboa, Centro de História da Universidade de Lisboa,
Alameda da Universidade, 1600-214, Lisboa, Portugal
anaroldao@edu.ulisboa.pt

** <https://orcid.org/0000-0002-5980-0075>

Universidade de Aveiro, Campus Universitário de Santiago, 3810-193, Aveiro, Portugal
jpsilvestre@ua.pt

*** <https://orcid.org/0000-0002-3206-6878>

Universidade de Zurique, Romanisches Seminar, Zurichbergstrasse 8, 8032, Zurich, Suíça
joana.faroserafim@uzh.ch

a comunidades do Reino de Portugal pelos monarcas desde o século XII até 1279, bem com todas as suas cópias até ao final do século XV. Desenvolvido por uma equipa de investigação pertencente às áreas da História, da Linguística e das Ciências Computacionais, este sistema de edição digital – EDICOLAB – caracteriza-se por ser uma instância de trabalho colaborativo e multidisciplinar. Essa circunstância promoveu recorrentes reflexões sobre diversos estádios da implementação de todo o sistema, assim como momentos de experimentação que conduziram a escolhas finais nem sempre coincidentes com as ideias iniciais da equipa. A escolha do tipo de edição, as opções de fixação do texto e do grau de intervenção do editor, a solução digital encontrada e as contingências tempo/financiamento são alguns dos aspetos em debate neste artigo.

PALAVRAS-CHAVE edição digital, forais portugueses, humanidades digitais

ABSTRACT This article describes the development, testing and implementation phases of EDICOLAB, a digital editing system adapted to a corpus of non-literary manuscripts, namely Portuguese municipal charters. The *corpus* consists of royal municipal charters issued to the communities of the Kingdom of Portugal from the 12th century to 1279, as well as their copies up to the end of the 15th century. Developed by a research team that spans History, Linguistics and Computer Science, EDICOLAB embodies a collaborative and multidisciplinary approach. This led to recurring reflections on various stages of the implementation of the whole system, as well as moments of experimentation that led to final choices that didn't always coincide with the team's initial ideas. This article addresses fundamental considerations such as the choice of editing methodology, text stabilisation options, levels of editorial intervention, the digital solutions adopted and the impact of time and financial constraints.

KEYWORDS digital edition, portuguese municipal charters, digital humanities

As mais antigas cartas de foral, concedidas por monarcas portugueses a comunidades do seu território, assumiram-se como o objeto de estudo de um projeto de investigação denominado *iForal; Forais Medievais Portugueses: uma perspectiva histórica e linguística na era digital*, sediado na Faculdade de Letras da Universidade de Lisboa.¹ Desenvolvido por uma vasta equipa das áreas de conhecimento da História, da Linguística e das Ciências Computacionais, pertencente a seis de universidades portuguesas e a uma universidade no estrangeiro, a Universidade de Lausanne, o *iForal* tem, por objetivo central, a edição, em formato digital, dos mais antigos forais régios portugueses até ao final do reinado de D. Afonso III (1279) e as cópias que deles foram feitas até ao final do século XV. Essa edição digital é ainda acompanhada de um glossário de termos portugueses constantes nesses manuscritos (Roldão; Serafim, 2021).

Num período de quase três séculos, as cartas de foral foram atribuídas às comunidades urbanas com fins eminentemente políticos, com vista à estabilização territorial, económica e social do Reino, tendo as comunidades passado a contar com um documento/monumento regulador das relações com o rei e também da sua vida coletiva (Coelho, 1996; Reis, 2002). A carta de foral interessava, então, a todos os intervenientes que nela participavam. Por esse motivo, os reis foram, sucessivamente, confirmando os privilégios e liberdades atribuídos às comunidades pelos seus antecessores, assim como as cidades procuraram preservar e atualizar os conteúdos normativos do foral, designadamente através de pedidos de clarificação dos significados do que ficava regulado. Na verdade, até ao ano de 1279, todas essas cartas teriam sido originalmente escritas em latim, e foram “vertidas” para linguagem vernacular ao longo do período medieval, não só no interior da chancelaria régia, mas, e sobretudo, nas escrivainhas das cidades, por escrivães e tabeliães ao serviço das assembleias municipais.

O presente artigo visa demonstrar as diferentes fases de conceção, experimentação e implementação de um sistema de edição digital, no seio de uma investigação de natureza interdisciplinar, aplicada a

1 O website oficial do projeto é o seguinte: <https://iforal.hypotheses.org/>.

um *corpus* de manuscritos não literários: as cartas de foral. Na primeira parte, procuraremos abordar algumas das principais questões que ensaiámos responder ao longo da investigação, no que concerne ao tipo de edição a adotar de acordo com as características internas e externas do *corpus*, assim como com os objetivos da edição e, ainda, as estratégias de fixação de texto e de uso da informação contida na documentação. Na segunda parte, pretendemos descrever a solução digital proposta pelo projeto para a edição da documentação, o sistema de edição EDICOLAB.

Mergulhemos, então, nas indagações e nos debates que, durante os últimos quatro anos, paleógrafos, diplomatas, historiadores, filólogos, editores de outros textos, linguistas e informáticos, da equipa do projeto *iForal*, têm promovido sobre edição em formato digital.

PROBLEMATIZAÇÃO, DEBATE E EXPERIMENTAÇÃO

A primeira reflexão da equipa consistiu em pensar como, para quem e com que objetivos se pretendia editar os forais régios. Se era claro que a edição documental deveria satisfazer, por um lado, critérios científicos de produção de conhecimento heurístico sobre essa fonte documental, potenciando o seu estudo junto da comunidade académica e científica, por outro, era igualmente claro que o impacto societal de disponibilizar às comunidades urbanas conhecimento sobre um património coletivo tão valorizado e comemorado ao longo dos tempos seria uma missão igualmente fundamental (Roldão; Serafim, 2021). Nesse sentido, a edição não poderia ser apenas para especialistas, com fins exclusivamente científicos, mas deveria ser, também, dirigida ao público em geral, contribuindo para a rememoração de um passado identitário, estruturante e coletivo no interior da própria comunidade, por meio de iniciativas culturais e pedagógicas promovidas por câmaras municipais, museus, bibliotecas, escolas e sociedade civil.

Pensar em que edição queríamos para essa fonte, de acordo com esses objetivos, fez-nos refletir, seguidamente, sobre o próprio ato editar, o seu significado, as competências que envolveria e as formas que

poderia alcançar. Editar, diríamos, é sinónimo de identificar e estabelecer letras e traços que o editor transporá, com uma adaptação variável, do suporte original do texto manuscrito para o suporte onde figurará o texto editado. Nesse momento, cristaliza-se o estado, o conteúdo e a forma do texto, inaugurando-se uma nova fase da sua sobrevivência: o texto manuscrito passa a conservar-se também fora do suporte para o qual foi originalmente concebido. Esse texto sobrevive em uma dupla vida, no arquivo e fora dele, adquirindo uma existência mais complexa, desejavelmente mais utilizada, mais escrutinada e mais indagada. Contudo, essa condição não o torna imune ao risco de desaparecimento, quer no arquivo, quer mesmo em formatos de edição digital que tendem a transmitir maior certeza de comunhão e perenidade da informação, por contraponto à edição em papel. Em qualquer dos casos, o manuscrito continua a ser um monumento a preservar, nos vários suportes em que sobrevive, prosseguindo como artefacto sujeito à passagem do tempo e aos seus intervenientes. Editar é, assim, o resultado de competências especializadas, como o estabelecimento crítico de formas utilizadas, a interpretação de significados e de comparação crítica de textos com mais de um testemunho e, por vezes, o domínio da leitura paleográfica, como é o caso dos textos medievais. O editor é sempre interveniente, ainda que em uma ampla gradação de possibilidades (Blecua, 1990; Faulhaber, 1991).

Os manuscritos que aqui nos ocupam, as cartas de foral, evidenciam características internas e externas desafiantes para um editor. A informação jurídico-histórica que veicula e o seu carácter não literário têm vindo a colocá-lo tradicionalmente sob os olhos do historiador, e menos sob os olhos do linguista. Talvez por isso os episódios de edição de forais que conhecemos estão, sobretudo (ainda que não exclusivamente), na alçada de paleógrafos, diplomatas e historiadores, desde o século XIX. Contudo, o facto de terem conhecimento das cartas de foral ou textos de natureza semelhante para períodos tão recuados como o século XII (ou até antes), com múltiplos testemunhos ao longo de todo o período medieval, abarcando, assim, diferentes estádios da língua escrita em território português, naturalmente atraiu e atrai a atenção de linguistas e filólogos.

Estará o editor de cartas de foral preparado para fixar um texto capaz de atender, simultaneamente, àquilo a que comumente se denomina de edição para historiadores e de edição para linguistas? Ou, dito de outro modo, terá o editor competência para editar um texto de modo a ser utilizado por historiadores e linguistas, não perdendo dados pertinentes para o grau de indagação especializada que cada um deles procura nos seus respetivos domínios científicos? Nesse, como noutros casos, a opção por uma edição desenvolvida em ambiente colaborativo interdisciplinar parece ser o caminho mais razoável, em linha com as práticas de investigação de tantas outras áreas do conhecimento. Contudo, o reverso dessa comunhão é encontrar a convergência equilibrada de soluções para a edição, sempre num intervalo de tempo que é limitado à vigência de financiamento dos próprios projetos de edição.

Definido o caminho colaborativo e interdisciplinar, a opção pela edição digital tornou-se evidente, consumando-se logo um contributo da crítica textual para os domínios do conhecimento histórico: o uso de ferramentas de edição em ambiente digital aplicadas ao texto medieval não literário. Chegados aqui, impõe-se responder à seguinte questão: que desafios colocou o *corpus* documental do projeto ao editor?

O trabalho de pesquisa documental e crítica dos mais antigos forais régios portugueses revelou a sobrevivência de cerca de 400 testemunhos de cartas de foral, atribuídos a um total de cerca de 180 comunidades urbanas, em um período que se prolonga por quase três séculos. Ainda que, em cada caso, os números precisos possam variar, a verdade é que, em média, cada localidade terá conhecido, pelo menos, dois a três testemunhos do seu foral, sendo escasso o número de terras apenas com um único testemunho sobrevivente. Qualquer que tenha sido a situação precisa, verificámos, também, que só em 5% dos casos sobreviveu o documento original. O apuramento desses números permite afirmar que estamos, maioritariamente, perante textos com múltiplos testemunhos, com perda do documento original mais antigo, o que de imediato nos interroga sobre o tipo de edição mais adequado. Com efeito, devemos editar, um a um, cada testemunho sobrevivente de um foral, ou procurar estabelecer um texto crítico (ou fixar o mais

antigo testemunho) que albergue as variações constantes dos restantes testemunhos? Essa indagação não é, como seria de esperar, exclusiva do caso português, e coloca-se recorrentemente, por exemplo, aos editores de *fueros* dos restantes reinos peninsulares no período medieval (Castillo Llach, 2024). A opção de edição do *iForal* recaiu sobre a primeira hipótese, e a justificação assenta na diversidade de informação que cada testemunho de um mesmo foral encerra, sob diferentes pontos de vista.

Com efeito, de testemunho para testemunho de uma mesma carta de foral, a reescrita da informação poderá apresentar circunstâncias de produção muito diferenciadas. Por um lado, podemos estar perante manuscritos emanados de uma mesma instância de produção, como a chancelaria régia, sob a forma de documentos avulsos ou registados em diferentes livros de registo de um mesmo monarca ou mesmo de monarcas diversos que confirmam *in extensum* as cartas de foral outorgadas pelos seus antecessores. Por outro lado, podemos detetar instâncias distintas de produção dessas cartas, como, por exemplo, documentos saídos da chancelaria régia e documentos produzidos no âmbito das chancelarias urbanas. Ainda que possamos prosseguir na casuística dos dados, a verdade é que, no interior da tradição manuscrita de uma mesma carta de foral, poderá coexistir e, de certo, coexistirá variação das autorias jurídica, diplomática e material (as últimas, quando identificadas), assim com variação na data cronológica e tópica, na *mise en page* da informação e do uso de sinais de validação, e no grau de solenidade do manuscrito, por exemplo.

Do mesmo modo, e para além desses aspetos, a repetição da informação, sob a forma de reescrita por cópia simples ou em pública forma de um testemunho, eventualmente perdido ou ainda não identificado, terá introduzido não só variantes de escrita importantes, como as que encerram valor paleográfico, mas, sobretudo variantes significativas do ponto de vista semântico, com valor filológico e linguístico. A adoção pelo latim como língua de escrita ou a opção pela redação em vernáculo poderá, igualmente, estar presente na tradição manuscrita de uma mesma carta de foral. Cada conjunto de testemunhos atribuídos a um mesmo foral parece, em suma, poder encerrar uma ampla paleta

de circunstâncias e de sentidos que dificilmente resultaria perceptível e funcional para o trabalho concomitante de historiadores e linguistas numa edição crítica de texto com apresentação de variantes. O projeto *iForal* apresenta a edição de cada testemunho de um mesmo foral, a qual se seguirá o processo de colação desses testemunhos, feita com o auxílio de ferramentas de colação automatizadas.

Ainda como justificação para o tipo de edição seguido, devemos referir, por fim, às relações de dependência que essas cartas de foral estabelecem entre si, formando o que hoje se conhece como famílias e sub-famílias de forais (Roldão; Serafim, 2021). Ora, para corroborar a identificação e composição das famílias já existentes ou para as reelaborar, é essencial a análise crítica não só da informação contida nos diferentes testemunhos de um mesmo foral – única abordagem até agora seguida –, como também dos dados de natureza linguística que cada um apresenta. Essa análise permitirá identificar, por exemplo, o testemunho de uma carta de foral-modelo para a redação de uma outra carta dela resultante.

Uma vez definido o tipo de edição a desenvolver, a equipa confrontou-se com a tarefa de fixação do texto, o que necessariamente implica escolher critérios de transcrição a adotar consistentemente. Os critérios de transcrição constituem sempre um ponto de reflexão importante quando pretendemos editar um documento (Castro; Ramos, 1986; Costa, 1993). Os critérios a adotar variam muito do tipo de publicação ou edição que se pretende, do seu objetivo e, ainda, do público-alvo. No caso do projeto *iForal*, há os públicos-alvo bastante distintos: historiadores, linguistas e a comunidade, ou seja, todos os interessados em conhecer o foral da sua cidade, mas que não são académicos. Assim sendo, que critérios adotar e como fazer uma edição que sirva interesses tão distintos?

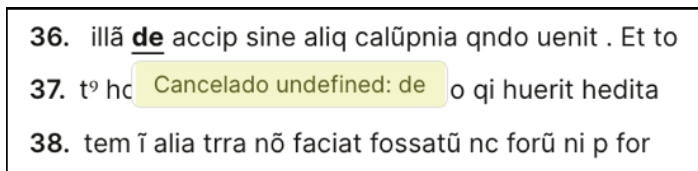
Para conseguir responder a essa pergunta, estabelecemos, como objetivo principal, dar a conhecer o texto tal como ele foi escrito, mas, ao mesmo tempo, de forma que fosse inteligível. Para isso, em primeiro lugar, definimos que a transcrição seria bastante conservadora, introduzindo apenas maiúsculas nos nomes próprios (topónimos e

antropónimos) e minúsculas nos nomes comuns, fazendo a separação de palavras e colocando pontuação estritamente nos casos em que se revela indispensável; para os textos latinos, procedemos também a uma normalização da oscilação da grafia das letras *u/v* e *i/j* para *i* e *u*, e respectivas maiúsculas *I* e *V*, independentemente do seu valor vocálico ou consonântico. E o que fazer em relação às abreviaturas? Se, por um lado, o seu desenvolvimento constitui uma intervenção no texto, por outro lado, um texto com palavras abreviadas não é inteligível.

Para respondermos a essa questão, criámos dois tipos de edição: uma edição paleográfica e uma edição crítica. A edição paleográfica não pretende ser uma reprodução fiel dos desenhos das letras e respetivas abreviaturas, mas inclui apenas as letras que estão representadas no manuscrito e alguns sinais, como o *et* tironiano, a abreviatura *us* (9, que geralmente surge na entrelinha superior) e a abreviatura *com-/ con-* (9).

As restantes abreviaturas não se encontram representadas na nossa edição paleográfica, pois percebemos que, pela enorme quantidade de palavras abreviadas dos nossos textos (cada palavra tem quase sempre pelo menos uma abreviatura, sendo frequente ter duas ou até três) (Roldão; Serafim, 2024), a responsividade do sistema de informação seria comprometida. Disso também não resultaria uma grande vantagem, uma vez que as edições têm sempre a imagem do manuscrito ao lado. Assim, optámos por manter apenas os referidos três sinais. A edição paleográfica conserva também os erros (repetições, por exemplo) e as emendas do copista (ex.: cancelamentos, palavras ou letras acrescentadas num segundo momento da escrita), identificados com “negrito” e visíveis numa caixa de texto (Ver Figura 1).

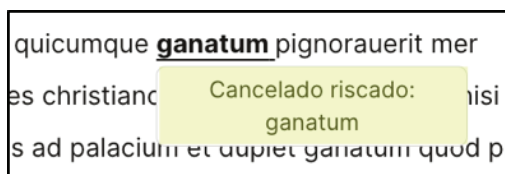
Figura 1: Identificação de palavra cancelada (edição paleográfica)



Fonte: Foral de Castelo Mendo (1281), EDICOLAB

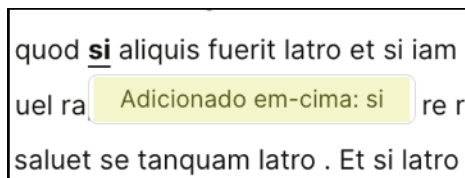
A edição crítica, pelo contrário, tem uma maior intervenção do editor, embora se mantenha o duplo objetivo de reproduzir o texto o mais fielmente possível, mas também de forma inteligível para os diferentes públicos. Assim, na edição crítica, as abreviaturas são desenvolvidas e o texto é reconstituído de acordo com a sua última versão, ou seja, privilegiam-se as formas corrigidas pelo copista, ficando as palavras ou letras canceladas visíveis ao selecionar-se com o cursor (Ver Figura 2). Nessa edição, os erros evidentes do copista são corrigidos (por exemplo, as repetições são anuladas; não se corrigem questões sintáticas e oscilações gráficas) e os acidentes do texto (adições, por exemplo) são registados, marcando em negrito a intervenção do editor e, ao selecionar a palavra com o cursor, é possível visualizar a versão que se encontra no manuscrito (Ver Figura 3).

Figura 2: Identificação de palavra cancelada (edição crítica)



Fonte: Foral de Coruche (1218), EDICOLAB

Figura 3: Identificação de palavra adicionada pelo copista (edição crítica)



Fonte: Foral de Évora (1218), EDICOLAB

Por outro lado, considerámos que a materialidade do texto também deveria estar presente e, por isso, as nossas transcrições reproduzem a disposição do texto original, registando-se as mudanças de linha, fólio e coluna, mas também as mudanças de mão e a existência

de sinal rodado; os títulos de cada parágrafo, geralmente destacados no texto original, são assinalados em negrito e, sempre que há rubricas, introduz-se parágrafo e conservam-se os caldeirões presentes do texto, recorrendo ao respetivo sinal gráfico. No que diz respeito às notas marginais, as anotações coevas nas margens de cada fólio são transcritas em nota e as anotações comprovadamente posteriores à produção do documento são ignoradas, exceto em situações que se considerem relevantes para o texto; nesse caso, são também transcritas em nota de rodapé, referindo-se “anotação tardia”.

Acreditamos que essas soluções são adequadas, que permitem ter, de forma fácil e intuitiva, acesso ao texto original, mas também a uma versão de texto desenvolvido e editado, satisfazendo as necessidades de historiadores, linguistas e do público em geral. É importante referir, ainda, que, em projetos que venham a integrar a plataforma desenvolvida pelo iForal, o sistema de edição não obriga, no entanto, a um formato único de edição e muito menos a determinados critérios de transcrição. Cada investigador ou grupo de trabalho poderá sempre adotar os seus critérios e optar pelo tipo de edição que se adequa aos seus objetivos.

Finalmente, o elenco da informação a fornecer sobre cada testemunho foi, igualmente, um desafio para a equipa. Uma das maiores dificuldades prende-se com a multiplicidade de informações contidas nos manuscritos sobre as circunstâncias de produção desses mesmos textos, às quais já aludimos acima. Repare, por exemplo, no caso da datação. Com efeito, um testemunho não original de uma carta de foral poderá indicar, direta ou indiretamente, a data da primeira concessão régia da carta à comunidade a que se destinou e, em simultâneo, como seria de esperar, a data em que esse testemunho foi produzido. Isso também se aplica quando o testemunho foi emanado por um monarca posterior à concessão, por autores de natureza diplomática e material distintas. Além disso, se um testemunho de uma carta de foral afirma explicitamente que segue, no seu articulado, uma outra carta de foral, essa informação é de todo imperdível. Ao fim e ao cabo, cada testemunho pode conter a referência a diferentes camadas de tempo, diferentes lugares de produção,

autores e intervenientes (como testemunhas e confirmantes), e até outros testemunhos, cuja enunciação deve, segundo o nosso entendimento, constar dos metadados apresentados para cada um deles.

Ainda no que concerne à informação que os forais apresentam, destaca-se um último desafio lançado ao editor, sobretudo ao editor de uma edição em formato digital. As disposições normativas contidas nos forais sobre o funcionamento das comunidades e a sua relação com outros poderes instituídos, nomeadamente, o poder régio, promoveu o registo de informação muito variada e, em alguns casos, detalhada sobre áreas fundamentais da vida municipal como as atividades económicas, a fiscalidade, a organização social, a justiça, entre outras. Referências a ofícios e cargos, a grupos sociais, a bens de consumo e/ou de troca comercial, a moedas, a penas, a impostos, entre tantas outras, assim como a enunciação de nomes próprios e de lugares, são comuns a todas as cartas de foral. As edições em papel tendem, naturalmente, a elencar essa informação em índices especializados, designadamente em índices antroponímicos, toponímicos e ideográficos. Contudo, a edição digital poderá apresentar outras formas de visualização da informação. A opção seguida no projeto *iForal* consistiu na atribuição ao leitor da tarefa de gerar as informações que procura no interior das funcionalidades do próprio sistema integrado de edição, designadamente através de listas alfabeticamente ordenadas de todas as palavras registradas na documentação e que geram níveis de concordância pesquisáveis. Ademais, está assegurada a ligação a um glossário de termos portugueses, onde se explicitam significados e ocorrências no interior de todo o *corpus*. Na verdade, o amplo alcance da pesquisa eletrónica de texto, a gestão de tempo de trabalho dos editores, assim como o incalculável e infundável elenco dos interesses específicos dos usuários do sistema parece justificar essa escolha.

As decisões tomadas sobre como editar forais régios medievais resultaram de um amplo e longo diálogo interdisciplinar entre os membros da equipa do projeto. Em periódicas reuniões de trabalho, historiadores e paleógrafos com experiência no estudo e edição de documentação foralenga, por um lado, e linguistas especializados em

história da língua portuguesa, em crítica textual e em edições eletrônicas, por outro, cedo perceberam a necessidade de estabelecer com clareza os conceitos que, em cada área de trabalho, eram utilizados. O mais relevante de todos foi, sem dúvida, o conceito de edição eletrônica, de linguagem de marcação e os métodos a ela associados. Tratando-se de um campo quase desconhecido para os historiadores, antes familiarizados com a edição de texto em papel, foi fundamental a promoção de sessões práticas de trabalho sobre os tipos de edição eletrônica mais utilizados e sobre a natureza da linguagem XML-TEI e os respetivos procedimentos de implementação.

Ainda no domínio dos conceitos, foi necessário igualmente clarificar as nomenclaturas associadas à fixação de texto e identificação das diferentes autorias, intervenções presentes num manuscrito e acidentes de texto, para que todas elas ficassem esclarecidas, quer para historiadores, quer para linguistas (por exemplo, autor diplomático, correção ou conjectura, entre outros conceitos essenciais à construção da plataforma). Nesse sentido, os menus de preenchimento da plataforma (informações sobre cada manuscrito e marcação das particularidades de cada texto) foram parametrizados durante meses, em estreita colaboração com os colegas da equipa das áreas das tecnologias de informação.

No domínio interpretativo, a experiência da interdisciplinaridade permitiu que linguistas e historiadores recentrassem a sua atenção, em conjunto, nas palavras que compõem cada texto.

Foi particularmente relevante a reflexão sobre vocábulos fundamentais, que ocorrem com muita frequência em forais medievais e que, nessa circunstância, puderam ser observados por especialistas com conhecimentos não necessariamente coincidentes, mas complementares, nomeadamente ao nível do significado histórico e diacrónico de termos relacionados com nomes de ofícios e de cargos no domínio económico e político, assim como de tributos, de grupos sociais, entre tantos outros; também ao nível da evolução da língua, seja no plano lexical, sintático ou morfológico. Com efeito, historiadores e linguistas ampliaram o conhecimento sobre cada palavra do manuscrito, promovendo um panorama interpretativo dos textos muito mais diversificado e inovador.

EDICOLAB, UM SISTEMA DE EDIÇÃO DIGITAL COLABORATIVA

O Sistema de Edição Digital Colaborativa (EDICOLAB, 2023) foi desenvolvido com o objetivo de apoiar e simplificar a colaboração no processo de preparação de uma edição crítica, quando realizada por uma equipa multidisciplinar, permitindo também a publicação numa página web acessível ao público.

A ferramenta EDICOLAB pode ser comparada a alguns projetos de edição digital, como o CEED (*Cooperative Web-Based Editor for Critical Editions*), o *Patrimonivm Editor* e o EVT (*Edition Visualization Technology*). O CEED, desenvolvido no contexto do projeto PhiBor, tem como objetivo facilitar a edição colaborativa de manuscritos medievais, permitindo a divisão dos documentos entre diferentes utilizadores. Embora seja uma solução bastante robusta, não inclui um glossário e apresenta limitações no que toca ao visualizador de documentos, tornando-o insuficiente para projetos que necessitam dessas funcionalidades adicionais (Marotta, 2021). O *Patrimonivm Editor*, por sua vez, destaca-se pela funcionalidade de georreferenciação dos locais mencionados nos manuscritos, sendo útil para projetos multidisciplinares, mas também não inclui um glossário, focando-se em necessidades mais específicas (Broux, 2019). Já o EVT é uma aplicação voltada principalmente para a visualização de edições críticas em TEI e simplifica a apresentação dessas edições ao público. Contudo, não suporta edição colaborativa nem inclui um glossário, limitando-se a facilitar a visualização de edições críticas (Rosselli Del Turco, 2019).

O investimento numa ferramenta própria, como o EDICOLAB, justifica-se pelas limitações das soluções existentes, que não atendem plenamente às necessidades de um projeto de edição crítica colaborativa, especialmente quando envolve equipas multidisciplinares. Uma das razões principais para o desenvolvimento do EDICOLAB é o facto de muitas ferramentas exigirem conhecimentos especializados em linguagens de marcação, como o XML TEI, especialmente quando essa edição envolve um aparato crítico.

Um dos principais requisitos, no que diz respeito ao apoio ao trabalho editorial, é possibilitar que os membros da equipa responsáveis por cada tipo de informação possam registar os dados de forma independente, contribuindo diretamente para a edição digital. A implementação de processos de registo direto também permite diminuir o número de membros da equipa exclusivamente dedicado à codificação de dados elaborados por outros. A gestão de permissões é uma funcionalidade essencial em plataformas de edição colaborativa, pois é necessário regulamentar as ações que cada utilizador pode realizar num documento. Geralmente, esse controlo é atribuído com base em diferentes funções definidas pelo administrador do sistema. No caso do EDICOLAB, há três funções distintas associadas a um documento. Uma concede apenas permissões de leitura, outra permite tanto leitura quanto escrita, enquanto a terceira corresponde ao criador do documento. Este tem a capacidade de adicionar e remover utilizadores dos documentos que criou.

Para cumprir o objetivo de criar um ambiente colaborativo, procurou-se implementar um sistema em que vários utilizadores possam contribuir para a construção da edição de forma simultânea. Os administradores deverão ser capazes de gerir as permissões de leitura e escrita dos diferentes utilizadores no modo de edição, garantindo um controlo eficaz sobre quem pode aceder e modificar o documento. Por fim, para facilitar o processo de edição para utilizadores com diferentes níveis de experiência, procurou-se implementar um visualizador TEI que construa uma representação visual a partir das etiquetas XML e que valide o código-fonte a nível sintático para identificar erros de marcação.

A edição colaborativa é uma funcionalidade associada a uma aplicação de edição de texto, que permite que dois ou mais utilizadores modifiquem o mesmo documento. Esse processo pode ocorrer de forma concorrente ou não concorrente. No primeiro cenário, todos os utilizadores ligados a um documento podem realizar alterações simultaneamente, sem interferir no trabalho dos outros. Isso é possível porque as alterações de cada utilizador são feitas em uma cópia local do

documento, que é posteriormente sincronizada de forma assíncrona com os seus pares.

No segundo cenário, os editores podem modificar o mesmo documento, mas não ao mesmo tempo. É usado um mecanismo de bloqueio (*locking*) que garante que, quando um utilizador abre um documento para edição, ele tem exclusividade para o modificar. Somente quando o utilizador fecha o documento é que outro pode assumir o controlo e fazer alterações. Implementar o primeiro cenário é mais complexo do ponto de vista da programação e consome mais recursos de alojamento, pelo que é conveniente avaliar a real necessidade dessa funcionalidade. No do projeto *iForal*, e tendo em conta a organização do trabalho de edição em que raramente intervêm simultaneamente vários utilizadores, concluiu-se que não se justificaria o investimento numa funcionalidade tão complexa (Sousa, 2023, p. 9).

No que diz respeito à edição do texto, destaca-se o desenvolvimento de um editor de texto WYSIWYG, que permite a marcação direta do corpo do documento. O editor é uma componente da aplicação em que os utilizadores podem efetuar marcações TEI de forma flexível ou restrita. Na marcação flexível, é possível escrever as etiquetas XML diretamente numa caixa de texto, sendo mais recomendada para utilizadores experientes que desejam manipular o código-fonte TEI de forma direta. Na marcação restrita, as marcações são feitas através de menus, botões ou outros elementos que fornecem nomes amigáveis às etiquetas TEI, sendo indicada para utilizadores familiarizados com os termos técnicos da edição crítica, mas não necessariamente conhecedores da codificação TEI.

Além da forma como as marcações são registadas, as aplicações de edição também podem apresentar o código-fonte de duas maneiras distintas. Num dos paradigmas, utiliza-se o princípio WYSIWYG, no qual o código-fonte é ocultado, apresentando-se, em vez disso, uma pré-visualização da interpretação gráfica das etiquetas XML. Outra possibilidade é o princípio “*what you see is what you mean*” (WYSIWYM), em que o código-fonte é apresentado diretamente ao utilizador. A marcação restrita do código-fonte é assegurada pela presença de menus (Ver

Figura 4), que facilitam a personalização de etiquetas TEI para utilizadores menos experientes. A definição do número e estrutura dos menus tem em conta a previsibilidade da marcação dos textos, com operações. Considerando que o conjunto de etiquetas utilizadas nesse projeto tem uma aplicação previsível e recorrente, foi desenvolvido um menu de marcação que aplica sistematicamente as marcações necessárias para a descrição de dados de crítica textual (Sousa, 2023, p. 74).

Figura 4 – Menus de apoio à edição



Fonte: EDICOLAB

O módulo de edição apresenta, atualmente, menus que respondem à descrição de elementos estruturais, fenómenos do manuscrito e intervenções do editor. Os seis menus são: *Estrutura*, *Acidentes*, *Substituição*, *Editor*, *Abreviatura* e *Nota*. O objetivo é incluir, nos menus, todas as etiquetas necessárias ao processo de marcação TEI XML do corpus documental do projeto.

Ao menu *Estrutura* correspondem todas as etiquetas usadas para marcar questões estruturais, estando nele contidas oito etiquetas. As duas primeiras etiquetas, <text> - *Texto* e <body> - *Corpo*, são obrigatórias a todos os documentos XML TEI. A etiqueta <p> - *Parágrafo* é também indispensável nessa edição, uma vez que é obrigatório que as frases estejam inseridas num elemento <p> ou equiparado. Na fase inicial de desenvolvimento, em *estrutura* encontravam-se os três elementos

anteriormente referidos, e *Início de Página* (<pb>) e *Início de Linha* (<lb>), escolhidos por se considerar relevante marcar quebras de linha e de página. Após a análise de alguns documentos, essas etiquetas revelaram-se insuficientes para codificar os manuscritos em mãos. Uma das primeiras etiquetas a ser acrescentada a essa categoria foi <cb> - *Início de Coluna*, para evidenciar, na edição digital, a estrutura do documento físico cujo texto se reproduz. Respeitando a conservação da estrutura do manuscrito, implementou-se, na edição digital, a marcação das diferentes secções em que esse tipo de documento se divide. Existem pelo menos três grandes secções distintas – a secção legislativa, a confirmação do rei e as testemunhas – pelo que se acrescentou a etiqueta <div> - *Divisão*.

Os documentos do *Text Encoding Initiative* não contemplam um módulo destinado à codificação de Forais, por isso, a alternativa encontrada pela equipa foi o uso desta etiqueta <div>, junto com o atributo @type, para identificar e especificar as partes distintas do documento. Surgiu, ainda, a necessidade de acrescentar o elemento <ab> - *Bloco* à categoria, usado nessa edição para marcar as diferentes testemunhas quando surgem listadas individualmente em colunas no final do documento.

Ao menu *Acidentes* pertencem todas as etiquetas usadas para codificar acidentes materiais, problemas de leitura resultantes desses acidentes e ações do copista. No início, o menu contava apenas com três etiquetas, que, mais tarde, revelaram-se insuficientes para dar conta de todos os fenómenos presentes nos diferentes manuscritos. As etiquetas <add> - *Adição*, - *Cancelamento* e <unclear> - *Dúvida* foram as primeiras a configurar nesse menu e são usadas para marcar adições, cancelamentos ou porções de texto de leitura duvidosa, respetivamente. Juntaram-se às etiquetas <gap> - *Lacuna* e <damage> - *Acidente Material*, uma vez que o nível de degradação de alguns documentos era bastante elevado, chegando mesmo a provocar lacunas no texto. Assim, percebeu-se que seria necessário codificar os acidentes materiais que interferem diretamente no conteúdo do documento, usando a etiqueta <gap> para dar conta de lacunas materiais e a etiqueta <damage> para dar conta de acidentes do suporte.

O menu seguinte contém, também, etiquetas usadas para codificar ações do copista. No entanto, a necessidade de dar a entender ao utilizador não especializado que esses elementos devem ocorrer em conjunto, seguindo as regras da própria linguagem de marcação para codificar um determinado fenómeno, levou a figurarem um menu próprio, denominado *Substituição*. Nele estão contidas as etiquetas <subst> - *Substituição*, <add> - *Adição* e - *Cancelamento*, usadas em conjunto para codificar uma correção feita por meio de uma substituição. Embora dois desses elementos constem também no grupo anterior e, em rigor, codifiquem o mesmo tipo de intervenção no texto, o facto de ser necessário usar sub-etiquetas dependentes da etiqueta <subst> para uma substituição ser codificada levou à criação de uma secção separada para agrupar esses elementos. Surgindo no menu com uma organização de forma hierárquica pela ordem em que devem ocorrer na marcação, essas etiquetas serão facilmente utilizadas pelo editor de forma correta para codificar a operação, mesmo tratando-se de um utilizador não especializado.

O mesmo acontece com o menu *Abreviatura*. Nele estão contidas as etiquetas <abbr> - *Abreviatura* e <expan> - *Expansão*, que codificam uma abreviatura e a forma plena, respetivamente, e a etiqueta <choice> - *Escolha*. Esta última funciona como <subst>, ou seja, como etiqueta “mãe” das outras duas. Usando as três em conjunto, o editor oferece as duas leituras possíveis para uma palavra. É essa marcação que permite oferecer aos leitores a opção por uma visualização do texto correspondente ao documento (de tipo paleográfico), ou por uma visualização com uma maior legibilidade (de tipo interpretativo). Novamente, surgindo em categoria própria e organizadas pela ordem que devem ser inseridas, as marcações são usadas pelo editor de forma correta.

No menu *Editor* constam as etiquetas necessárias para codificar as intervenções do editor no texto. Essa categoria começou por conter apenas três etiquetas, usadas para codificar erros evidentes no texto e a sua respetiva correção: <choice> - *Escolha*, <sic> - *Erro* e <corr> - *Correção*. Novamente, as duas últimas sub-etiquetas dependentes de <choice>, usadas em conjunto, permitirão que o leitor escolha a

variante que quer ler. Posteriormente, foram acrescentadas as etiquetas <supplied> - *Conjetura* e <surplus> - *Repetição*. Essas alterações surgiram da necessidade de dar resposta a problemas levantados por alguns documentos que continham lições ausentes ou repetidas. No primeiro caso, quando o grau de ilegibilidade ou degradação do suporte é elevado, o editor pode recorrer à etiqueta <supplied> para facultar o texto em falta. Essa reconstituição é feita usando outro testemunho do mesmo texto. No segundo caso, quando surgem no testemunho lições repetidas que sejam consideradas desnecessárias, são codificadas com a etiqueta <surplus>. Esta não identifica a passagem como erro, mas revela que o editor a considera irrelevante para a leitura do texto. O menu estava, inicialmente, identificado como *Erro*, uma vez que continha apenas as etiquetas usadas para codificar os erros evidentes. No entanto, após o acrescento das outras etiquetas, percebeu-se que o nome *Editor* seria mais apropriado e abrangente, uma vez que passou a tratar-se de um menu que reúne etiquetas usadas para codificar as intervenções do editor sobre o texto.

Ao último menu deu-se o nome *Nota*. Como o nome pode indicar, contém a etiqueta <note> - *Nota*. Esse marcador é usado sempre que o editor achar relevante fornecer esclarecimentos adicionais sobre o texto/a passagem que está a codificar. Tratando-se, novamente, de uma intervenção do editor sobre o texto, poderia também figurar no menu *Editor*. Uma vez que esse sistema permite a alteração e organização das categorias e etiquetas conforme as necessidades da equipa, numa revisão ulterior essa etiqueta poderá ser englobada no menu *Editor*.

O sistema EDICOLAB foi desenhado prevendo a adaptabilidade a diferentes contextos. Antecipando a eventualidade de o sistema ser usado em projetos com especificidades distintas, desenvolveu-se um sistema dinâmico de gestão de etiquetas. Esse sistema permite aos administradores a liberdade de adicionar, editar ou remover as etiquetas que estão disponíveis no menu de marcação. A interface dedicada à gestão de etiquetas foi projetada com um design intuitivo, facilitando aos administradores a visualização e gestão. Consiste fundamentalmente numa quadro que apresenta as informações básicas de cada etiqueta:

Quadro 1: Etiquetas TEI e nomenclaturas

Nome amigável	Etiqueta TEI	Categoria de menu
Abreviatura	abbr	Abreviatura
Escolha	choice	Abreviatura
Expansão	expan	Abreviatura
Acidente Material	damage	Acidentes
Adição	add	Acidentes
Cancelamento	del	Acidentes
Dúvida	unclear	Acidentes
Lacuna	gap	Acidentes
Conjetura	supplied	Editor
Correção do Editor	corr	Editor
Erro	sic	Editor
Escolha	choice	Editor
Repetição	surplus	Editor
Bloco	ab	Estrutura
Corpo	body	Estrutura
Divisão	div	Estrutura
Início de Coluna	cb	Estrutura
Início de Linha	lb	Estrutura
Início de Página	pb	Estrutura
Parágrafo	p	Estrutura
Texto	text	Estrutura
Nota	note	Nota
Adição	add	Substituição
Cancelamento	del	Substituição
Substituição	subst	Substituição

Fonte: EDICOLAB.

ETIQUETAS TEI NO MENU CONFIGURÁVEL DE EDIÇÃO

O administrador tem a possibilidade de editar as etiquetas definidas e adicionar novas. Uma das secções da área de administrador consiste em uma lista de etiquetas que podem ser editadas individualmente através de um formulário. Para cada etiqueta, há vários campos a serem preenchidos, que configuram a utilização e demonstram a integração da nova etiqueta com o *standard* TEI (Ver Figura 5).

Figura 5: Formulário de configuração da etiqueta <abbr>

O formulário de configuração da etiqueta <abbr> é apresentado numa interface limpa com campos de texto e uma caixa de seleção. Os campos são:

- Nome:** Um campo de texto contendo o valor "abbr".
- Nome Amigável:** Um campo de texto contendo o valor "Abreviatura".
- Categoria:** Um campo de texto contendo o valor "Abreviatura".
- Contém filho?** Uma caixa de seleção marcada com um ícone de marcação, seguida pelo texto "Um filho é um marcador que existe dentro de outro".
- Atributos:** Um campo de texto contendo o valor "n", e um botão de texto "cancelled, damaged".

Fonte: EDICOLAB

O formulário recorre a designações descritivas que procuram guiar utilizadores não especializados. *Nome Amigável* refere-se ao título ou designação amigável da etiqueta, que será exibido no menu de edição, facilitando a identificação pelos utilizadores. *Nome da Etiqueta em Marcação TEI* é o nome que a etiqueta terá quando aplicada na marcação TEI. Esse nome deve respeitar as normas e convenções da linguagem de marcação. *Categoria* determina a que conjunto ou grupo a nova etiqueta pertencerá. Essa categorização tem como propósito organizar, de forma lógica e intuitiva, as etiquetas no menu. *Etiquetas Filho* refere-se à hierarquização das etiquetas, metaforizada na relação pais-filhos, indicando se possui sub-etiquetas dependentes. Por fim, o campo *Atributos da Etiqueta* especifica e detalha as propriedades da etiqueta.

Através desse formulário, os administradores têm autonomia para expandir e personalizar o conjunto de etiquetas disponíveis, adaptando o sistema às necessidades e especificidades do projeto em questão. Após realizar as alterações necessárias, os usuários podem guardar as modificações, atualizando, assim, a etiqueta na base de dados. Em termos de usabilidade e experiência do utilizador, é vital que as ferramentas de edição e inserção sejam consistentes e intuitivas, reduzindo, portanto, a curva de aprendizagem e potenciais erros durante o processo de gestão das etiquetas.

CONCLUSÃO

O caminho percorrido no projeto *iForal* permitiu-nos refletir sobre os desafios de conceber uma edição digital a partir das experiências específicas de distintas áreas de trabalho, que utilizam o documento manuscrito na sua investigação. Do mesmo modo, permitiu-nos avaliar em que medida as especificidades de um determinado *corpus* documental poderão condicionar as opções tomadas no momento de edição, no que concerne, por exemplo, à tipologia, ao nível de intervenção do editor e às características da fixação do texto.

O sistema de edição manuscritos, EDICOLAB, concebido no âmbito deste projeto, permitiu a concretização de um trabalho colaborativo e multidisciplinar que não só é adaptável às necessidades específicas de marcação de texto, no seio de contextos documentais diversificados (não literários ou literários), como possibilita que investigadores em diferentes estádios de conhecimento de marcação em linguagem XML-TEI possam operar a plataforma com êxito. Essas características potenciam a sua função pedagógica em contextos de futuras investigações em edição digital de outros manuscritos.

AGRADECIMENTOS

Agradecemos à mestranda Catarina Coelho, Bolseira de Investigação, as informações prestadas sobre questões críticas de funcionalidade dos

menus da plataforma EDICOLAB, cujo desenvolvimento acompanhou entre 2022 e 2024, quer como utilizadora, quer como membro ativo do projeto.

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

- BLECUA, Alberto. *Manual de Crítica Textual*. Madrid: Editorial Castalia, 1990.
- CASTILLO LLUCH, Mónica. In: *Los fueros y sus ediciones. Fueros Medievales: Fuentes jurídicas medievales para historiadores de la lengua*. Disponível em: <https://people.unil.ch/monicacastillolluch/>. Acesso em: 31 mar. 2024.
- CASTRO, Ivo; RAMOS, Maria Ana. Estratégia e tática de transcrição. In: *Colloque Critique Textuelle Portugaise*, 1981, Paris. *Colloque...* Paris: Fondation Calouste Gulbenkian / Centre Culturel Portugais, 1986. p. 99-122.
- COELHO, Maria Helena da Cruz. Concelhos. In: SERRÃO, Joel; MARQUES, Antônio Henrique Rodrigo de Oliveira (Dir.). *Nova História de Portugal* (Vol. III). Lisboa: Editorial Presença, 1996. p. 554-584.
- COSTA, Avelino de Jesus da. *Normas Gerais de Transcrição e Publicação de Documentos e Textos Medievais e Modernos*. Coimbra: Faculdade de Letras da Universidade de Coimbra - Instituto de Paleografia e Diplomática, 1993.
- EDICOLAB. *Sistema de Edição Digital Colaborativa*. Universidade de Aveiro. Disponível em: <http://edicolab-iforal.ua.pt>. Acesso em: 31 mar. 2024.

- FAULHABER, Charles Bailey. Textual Criticism in the 21st Century. *Romance Philology*, v. 45, n. 1, p. 123-148, 1991.
- REIS, António Matos. *Origens dos municípios portugueses*. Lisboa: Livros Horizonte, 2002.
- ROLDÃO, Filipa; SERAFIM, Joana. L'évolution graphique des actes, du latin au portugais: un exemple, la charte de franchise de Lisbonne (1179) et sa tradition. In: *Bibliologia Series*, Brepols Publishers, 2024 [submetido].
- ROLDÃO, Filipa; SERAFIM, Joana. Os mais antigos forais régios portugueses: Uma proposta de estudo e de edição. In: PEÑÍN, Raquel Martínez; DOMÍNGUEZ, Gregoria Caverio (Coords.). *Poder y poderes en la Edad Media*: Monografía de la Sociedad Española de Estudios Medievales. Murcia: Sociedad Española de Estudios Medievales, 2021. p. 375-386. Disponível em: https://medievalistas.es/wp-content/uploads/2021/08/poder-y-poderes_SEEM-16-WEB.pdf. Acesso em: 7 out. 2024.
- SOUSA, José. *Sistema integrado de suporte à edição eletrónica de forais medievais portugueses e de desenvolvimento de um glossário*. Dissertação (Mestrado em Engenharia Informática) – Universidade de Aveiro, Aveiro, 2023.
- BROUX, Yanne. PATRIMONIVM: Geography and economy of the imperial properties in the Roman world. 2019. Disponível em: <https://patrimonium.huma-num.fr>. Acesso em: 7 out. 2024.
- MAROTTA, Daniele. Tracing a digital critical edition: The experience of the PhiBor project. *Umanistica Digitale*, v. 5, v. 10, p. 89-114, 2021.
- ROSSELLI DEL TURCO, Roberto. Designing an advanced software tool for Digital Scholarly Editions: The inception and development of EVT (Edition Visualization Technology). *Textual Cultures: Texts, Contexts, Interpretation*, v. 12, p. 111-191, 2019.
- ROSSELLI DEL TURCO, Roberto; PIETRO, Chiara Di; JULIA, Kenny; RAFFAELE, Masotti. EVT - Edition Visualization Technology. 2014. Disponível em: <https://www.dish.unito.it/en/projects/digital-philology/evt-edition-visualization-technology>. Acesso em: 7 out. 2024.

Recebido: 31 mar. 2024 / Revisto pelo autor: 7 out. 2024 / Aceito: 04 set. 2024

Editor responsável: Ely Bergo de Carvalho