

Tema 2 Bases de dados relacionais enquanto ferramenta de investigação em História

André Evangelista Marques

Universidade Nova de Lisboa, Faculdade de Ciências Sociais e Humanas /
Instituto de Estudos Medievais (Portugal)
andre.omarques@gmail.com

Gabriel David

Universidade do Porto, Faculdade de Engenharia / INESC Porto (Portugal)
gtd@fe.up.pt

RESUMO

Este artigo resulta de uma investigação de doutoramento em história medieval, dirigida ao estudo da paisagem e do povoamento no NO de Portugal, entre os séculos IX e XI. A ferramenta escolhida para assegurar o tratamento de um grande volume de dados e garantir a flexibilidade da sua análise foi uma base de dados relacional. Neste texto procura-se discutir o processo de conceção desta base de dados, apresentar resultados e retirar algumas conclusões sobre aspetos mais gerais implicados no método definido. Apresenta-se a organização da informação, através da versão final do modelo de dados e de uma explicação dos seus principais componentes. E sublinham-se alguns aspetos relacionados com a investigação histórica propriamente dita. Para além da organização dos dados, procurou-se encontrar processos claros para os comunicar. Mostram-se alguns dos interfaces resultantes, assim como exemplos de estatísticas obtidas a partir da base de dados. As unidades espaciais foram georreferenciadas sempre que possível, pelo que foi possível alimentar automaticamente a respetiva cartografia. Um resultado final é a integração da base de dados num repositório institucional, de forma a disponibilizá-la em livre-acesso à comunidade científica.

Palavras-chave: Repositórios de dados, investigação quantitativa, bases de dados nas ciências sociais, e-Science

ABSTRACT

The context of the paper is a PhD research in medieval history, focused on landscape and settlement in NW Portugal, from the 9th to the 11th century. To solve the problem of organizing a large volume of data and guarantee flexibility in analyzing it the tool chosen has been a relational database. This paper's purpose is to discuss the process of designing such a database, present its results and draw some conclusions on the generality of the method. The organization of the information will be presented through the final design of the database model and by an explanation of its main constituents. Some details pertaining to the historical research needs will be highlighted. Besides the organization of the data, a significant effort has been devoted to find clear ways to communicate it. Some of the resulting interfaces will be shown, as well as some example statistics already obtained from the database. The territorial entities have been geo-referenced wherever possible so it has been possible to automatically feed the corresponding cartography. A final result is the integration of the database in an institutional data repository to make it openly available to the research community.

Keywords: Data repositories, quantitative research, databases in social sciences, e-Science

1. CONTEXTO

A base de dados (BD) aqui apresentada foi concebida no quadro de uma dissertação de doutoramento em história medieval dedicada ao estudo da paisagem e do povoamento no território da diocese de Braga, norte de Portugal, entre os séculos IX e XI. O enquadramento desta investigação nas propostas de J. Á. García de Cortázar (2004) para o estudo da *organização social do espaço* no quadrante NO da Península Ibérica, entre os séculos VIII e XIII, explica desde logo o problema em análise. Partindo do pressuposto teórico de que a uma qualquer sociedade, enquanto estrutura social, corresponde um esquema de distribuição interna do poder que se projeta na atribuição e organização específicas do espaço que essa sociedade ocupa, García de Cortázar concebeu um edifício conceptual e metodológico que insiste no conhecimento do território físico sobre o qual os poderes sociais tomam decisões, com vista à produção de bens, ao enquadramento de pessoas e à difusão de modelos culturais (García de Cortázar, 1998, p.334).

Na impossibilidade de um estudo completo da organização social do espaço na região escolhida, esta investigação optou por centrar a análise na paisagem e no povoamento, entendidos como bases materiais dessa organização. Não era possível avançar para a análise dos processos de apropriação, organização e articulação do território sem antes atentar em duas questões principais, que vieram a corporizar o objeto do trabalho: (i) os modelos discursivos de base e as circunstâncias conjunturais de transmissão que moldaram a representação documental do espaço; e (ii) a morfologia propriamente dita do espaço, analisada dentro dos constrangimentos impostos pelo registo escrito ao estudo das diversas unidades espaciais cuja tipologia os redatores distinguiram nos seus textos.

Procurando responder a este conjunto de perspetivas de análise de um objeto que definimos como a representação documental de uma realidade material socialmente construída, a investigação marcou dois objetivos principais: (i) o levantamento sistemático das distribuições cronológica e espacial das menções documentais às diversas unidades de organização do espaço; e (ii) o estudo morfológico dos vários tipos de unidades e da sua evolução. Estes objetivos decorrem diretamente da metodologia concebida por García de Cortázar (1988; 1999). Destacam-se apenas pela tentativa de alargar o leque de unidades (e escalas) espaciais em análise, desde as mais pequenas parcelas agrárias até às grandes unidades de articulação política do território.

Sem prejuízo da natureza muito variável dos tipos de unidades definidos, foi definido um questionário de análise que contempla um conjunto amplo de variáveis, divididas por três apartados. No primeiro, cabem as distribuições cronológicas e espaciais do conjunto de unidades de cada tipo. E utilizamos o plural na medida em que a análise dessas distribuições deve ter em conta diferentes cortes cronológicos e escalas espaciais: (i) a longa duração e a escala regional, que permitem relacionar a distribuição global dessas unidades com fatores propriamente geográficos e com a configuração estrutural do povoamento na região, o que explica em larga medida as manchas de maior e menor concentração; e (ii) o tempo curto e a escala micro-regional (ou mesmo local), que permitem relacionar a distri-

buição conjuntural dessas unidades com o duplo processo de organização do espaço e de construção da respetiva memória documental.

No segundo apartado, cabe um amplo conjunto de variáveis relacionadas com a caracterização morfológica: (i) os elementos de designação das unidades de cada tipo, que dizem muito sobre a sua morfologia; (ii) a integração espacial dessas unidades, tanto do ponto de vista físico (em paisagens concretas) como social (no quadro de malhas territoriais, de cariz administrativo ou simplesmente de domínio); (iii) a estrutura interna das unidades, o que obriga a considerar o conjunto dos componentes que elas podem integrar (e não apenas do ponto de vista físico: por vezes a integração é estritamente funcional ou mesmo patrimonial), bem como as possibilidades de combinação entre os diferentes tipos de componentes e respetivos mecanismos de articulação.

Por fim, num terceiro apartado cabem duas variáveis associadas à «atribuição social» (Peña Bocos, 1995) das unidades: (i) a cadeia de transmissão da titularidade sobre as unidades espaciais e (ii) o conjunto de imposições que sobre elas recaiam.

A quantidade de dados requerida pelo método proposto dificilmente seria tratável em tempo útil segundo os métodos tradicionais de organização da informação científica. A complexidade das relações entre as entidades envolvidas também desaconselha o recurso a ferramentas informáticas com pouca expressividade, como ficheiros de texto ou mesmo folhas de cálculo. Para resolver o problema de organizar um grande volume de dados inter-relacionados e garantir flexibilidade na sua análise, a ferramenta escolhida foi uma BD relacional, a qual oferece um nível de representação de conhecimento adequado ao registo de factos complexos, incluindo os metadados relativos às fontes dos dados.

O objetivo deste artigo é discutir o processo de construção de uma BD adequada à investigação pretendida, apresentar os seus resultados e tirar algumas conclusões sobre a generalidade do método. A secção 2 apresenta as necessidades de informação derivadas do método utilizado na investigação e, de seguida, o processo seguido no desenvolvimento da BD. A secção 3 apresenta o projeto da BD através do modelo de dados e dos formulários principais de interação. Os resultados obtidos são apresentados na secção 4, reservando-se a secção 5 para as conclusões principais.

2. MÉTODO

2.1 NECESSIDADES DE INFORMAÇÃO DO INVESTIGADOR

Antes de passarmos à descrição da BD que aqui apresentamos, importa elencar três *regras de método* essenciais nas propostas de García de Cortázar (1998; 1999), a que a nossa investigação procurou obedecer: (i) a preocupação com a recolha exaustiva de dados (que recusa a mera ilustração); (ii) o recurso à cartografia, entendida como instrumento de análise e não como mero veículo de demonstração, assumindo um papel essencial na hora de interpretar as distribuições espaciais das unidades em análise, que devem por isso ser objeto de uma localização tão precisa quanto possível; (iii) e a avaliação rigorosa das fontes utilizadas, atenta a possíveis desproporções no tempo, no espaço e na proveniência.

Destas *regras* decorrem três objetivos que a BD terá de garantir do ponto de vista dos mecanismos de recolha e tratamento dos dados: (i) a exaustividade da análise, tanto no plano quantitativo como qualitativo, o que obriga à criação de ferramentas auxiliares de análise estatística e de indexação, respetivamente; (ii) a rigorosa georreferenciação de toda a informação espacial, com vista à sua integração num sistema de informação geográfica

(SIG) capaz de a cartografar e, mais do que isso, de sustentar a análise relacional dessa informação com todas as variáveis geográficas relevantes; (iii) a caracterização tão rigorosa quanto possível do corpus documental compulsado e a análise do léxico espacial documentado, com particular atenção à tripla inscrição (temporal, espacial e social) de cada vocábulo. A estes objetivos cabe apenas acrescentar: (iv) a preocupação com a integração dos dados para os quais a base foi desenhada, de proveniência escrita, com dados de proveniência outra (geográfica, cartográfica, arqueológica, etc.), por forma a situar a análise no quadro necessariamente interdisciplinar em que o estudo de uma realidade como o espaço há de ter lugar.

2.2 METODOLOGIA DE PROJETO DA BASE DE DADOS

O processo de desenvolvimento da BD foi, como habitualmente, determinado pelas condições específicas da situação e guiado pelas necessidades de informação do investigador. Contudo, é normal que, durante a investigação, as necessidades de informação evoluam com os resultados intermédios obtidos e com a emergência de novas perspetivas. Por esta razão, a especificação da BD seguiu o desenrolar do processo de investigação e a chave do sucesso do projeto consistiu num diálogo intenso entre o investigador e o projetista da BD.

Foi portanto com base numa perspetiva de dinamismo significativo nos requisitos da BD que se optou por uma abordagem iterativa do projeto, opção essa reforçada pelo carácter de protótipo do resultado pretendido. Cada iteração foi constituída por quatro etapas principais:

- (1) esforço concetual por parte do investigador no sentido de explicitar as necessidades de informação correntes do processo de investigação;
- (2) entrevista entre o investigador e o projetista explicando o domínio de investigação, a perceção corrente das necessidades de informação e os resultados esperados;
- (3) (re)projeto da BD e da interface e migração dos dados anteriormente recolhidos para o novo modelo;
- (4) teste/uso da nova versão para o registo de dados e para a extração de resultados.

Houve ainda um passo final na última iteração destinado a documentar a BD, no sentido de a tornar utilizável por outros investigadores, tanto ao nível dos dados primários como dos resultados processados.

A BD resultante foi um instrumento da etapa de análise de conteúdo de fontes documentais, integrada num método mais abrangente, que estruturou a investigação em História. Nesta etapa, o investigador analisou exaustivamente o conteúdo de toda a documentação compulsada. A análise visava identificar todas as unidades espaciais de estruturação e ocupação do território, as suas características e as relações entre elas. O registo e posterior interpretação de toda a informação recolhida, em especial pela complexidade de relações em causa, não eram compatíveis com um registo em papel ou em ferramentas informáticas com pouca expressividade em termos de modelação. O facto de se ter optado por uma BD relacional facilitou uma representação informática do conhecimento progressivamente extraído dos documentos e a produção de evidências a suportar as hipóteses de investigação, por exemplo sob a forma de mapas anotados ou pela obtenção de indicadores quantitativos. O método desenhado pelo investigador para os trabalhos de preparação da sua tese de doutoramento foi determinado pela viabilidade de construir a BD.

3. O PROJETO DA BASE DE DADOS

O objetivo da BD construída foi, como referido na secção 1, registar factos acerca de unidades espaciais num determinado contexto espaço-temporal. O método seguido foi iterativo e, portanto, os requisitos da BD foram sendo adicionados ao longo do processo. Nesta secção optamos por apresentar apenas a versão final dos requisitos e uma explicação detalhada do modelo elaborado, com o duplo propósito de documentar o resultado obtido e de fundamentar a extração de conclusões úteis para futuros trabalhos de natureza similar.

3.1 REQUISITOS FUNCIONAIS E NÃO FUNCIONAIS

As unidades espaciais, sendo as entidades centrais da investigação, não são em geral observáveis por método direto, mas apenas através de referências documentais. São não tanto um ponto de partida mas mais um ponto de chegada da pesquisa. Por isso, assume-se pouco relativamente a estas unidades: uma designação normativa, um tipo normativo e uma localização geográfica, a preencher progressivamente. O tipo normativo é um vocabulário controlado com uma classificação hierárquica de até três níveis. A localização geográfica das unidades espaciais é baseada nas divisões administrativas de 2011 (INE) tendo os níveis de distrito, concelho, freguesia e lugar. Todas as divisões são caracterizadas pelo código oficial e pela designação. Tanto os lugares como as freguesias são ainda associados a uma localização em coordenadas geográficas, tipicamente o centro do lugar ou da freguesia.

Os factos centrais a registar na BD são as menções documentais às unidades espaciais, enquanto elementos de caracterização dessas unidades. Estes elementos são portanto encarados como entidades em si mesmos, descritos por um conjunto rico de atributos resultado da análise documental e atribuídos às unidades espaciais. À imagem do que sucede com as unidades espaciais, os elementos também são classificados diretamente de acordo com a tipologia normativa. A atribuição às unidades espaciais pode ser inequívoca, quando a informação existente é suficientemente explícita, mas por vezes é criado um novo registo de unidade espacial para fornecer um referente ao novo elemento e, mais tarde, vem-se a identificar essa unidade espacial com uma outra e procede-se à consolidação das atribuições de elementos a unidades espaciais, eliminando a repetição.

As menções documentais incluem por vezes informação acerca de relações entre as diferentes unidades espaciais. Essas relações são categorizadas e, como as menções documentais são concretizadas em elementos de caracterização das unidades espaciais, é entre os elementos que as relações são estabelecidas e não diretamente entre as respetivas unidades. São ainda de prever relações diretas com a tipologia normativa, ao invés de com um elemento de um determinado tipo.

Os tributos e as transmissões de propriedade relativos a elementos são objeto de registos próprios.

Atendendo à relevância do registo das fontes documentais, todos os elementos são ligados ao respetivo documento. Um documento pode ser a fonte de múltiplos elementos mas cada elemento é obtido de um único documento. Menções à mesma unidade espacial em documentos diferentes são registadas em elementos diferentes. Os documentos são identificados do ponto de vista arquivístico pelo seu produtor, pelo fundo a que pertencem e respetiva cota, tipologia e sumário. Um especial cuidado é posto na datação do documento, a qual inclui data, intervalo de datas e categorização por séculos, meios séculos e quartos de século.

Os documentos foram, em muitos casos, objeto de publicação e essa informação é registrada, sendo possível contemplar múltiplas publicações do mesmo documento.

Para além dos requisitos funcionais sumariamente descritos acima, foram identificados alguns requisitos não funcionais.

O sistema de gestão de bases de dados a utilizar deveria ser de fácil manuseamento por parte do investigador.

O número de factos a registar é da ordem dos milhares, o que não deverá colocar problemas de desempenho complexos.

As interfaces de registo de dados deverão ser informativas o mais possível e promover a consistência dos dados.

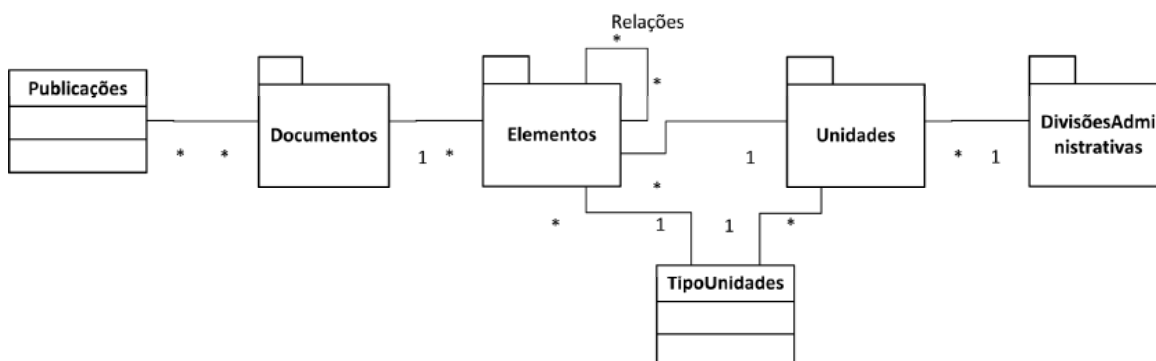
A extração de informação é de três tipos principais: listagens exaustivas dos registos, segundo várias perspetivas; indicadores estatísticos calculados segundo diversos critérios; e distribuições espaciais de parâmetros para representação cartográfica.

Como consequência dos requisitos não funcionais apresentados, selecionou-se o sistema de gestão de BD Microsoft Access 2007 para a implementação da BD, a que foi dada a designação «Paisagem e Povoamento (diocese de Braga, séculos IX-XI)».

3.2 MODELO DE DADOS

O modelo de dados final que derivou dos requisitos apresentados encontra-se esquematizado no diagrama UML de classes (Booch, Rumbaugh & Jacobson, 1999) simplificado da Figura 1. O modelo relacional detalhado pode ser consultado no Anexo A.

FIGURA 1: Esquema simplificado da BD



O módulo Documentos destina-se a registar os metadados relativos a todos os documentos consultados. Um documento pode estar publicado em várias publicações, as quais são descritas na classe Publicações. Por outro lado, cada documento pode conter vários elementos mas cada elemento pertence a apenas um documento. O módulo Elementos regista a informação relativa a cada menção documental, sendo que podem ser estabelecidas múltiplas relações entre dois elementos. Cada elemento é ainda atribuído a uma unidade, registada no módulo Unidades. Tanto as unidades como os elementos são classificados segundo uma tipologia de unidades. Finalmente, cada unidade é localizada numa divisão administrativa, desde o nível de lugar ao de distrito.

3.3 MODELO DE INTERFACE

A descrição detalhada dos atributos do modelo de dados é feita a partir da apresentação da interface constituída pelos formulários que correspondem aos três módulos principais

que estruturam a BD. No decurso desta descrição, agruparemos os diversos campos de cada formulário em secções temáticas que não ficaram explicitadas na BD por razões que se prendem estritamente com a visualização no espaço disponível das janelas. De qualquer forma, pareceu-nos útil reproduzir aqui uma imagem de cada formulário, para orientar a descrição e explicação do conjunto dos campos que os constituem.

A) Módulo Documentos

FIGURA 2: Formulário Documentos

The screenshot displays the 'Formulário Documentos' interface, divided into several sections:

- Top Section (A.1):** Contains fields for document identification: 'Nr' (843), 'Ano' (960), 'AnoF' (0), 'Mês' (7), 'Dia' (13), 'Data2', 'Séc.' (50), and 'Data2'. It also includes 'Fundo' (Livro de Mumadona Dias), 'Cota' (n.º 3), 'Crítica' (Autêntico), 'Tipologia' (Testamento), and 'Tradição' (Cópia de cartulário).
- Summary (A.1):** A text box containing a summary of the document's content, mentioning 'D.ª Chãmoa Rodrigues, devota, sentindo-se doente e depois de ser levada para o mosteiro de Guimarães...'
- Obs (A.1):** A text box with notes: 'TIPOLOGIA: «series notícia uel testamenti»', 'REFS. NÃO FICHADAS: Há neste documento várias unidades que não foi possível identificar e outras exteriores ao território da diocese de Braga que não foram fichadas...'
- Table (A.2):** A table listing elements with columns: 'Elemento', 'Tipo', 'Designação', 'Obj', and 'Unidade'. The table contains 13 rows of data, such as '3293 Villa uilla lalini' and '3294 Mosteiros locum monasterii vimaranes'.
- Bottom Section (A.1):** A text box containing a transcription note: 'Ct Doc - V. a transcrição anotada deste documento no Apêndice I.A. [Notícia] «In nomine domini. Flamula deuota filia Rudeici et Leodegundie dum uenit ad infirmitate timendo de extremo leuauit confessionem hic in uilla lalini...」'

Deste primeiro módulo constam, como ficou dito, dois tipos de informação: (A.1) os dados imprescindíveis a uma correta identificação e caracterização diplomática de cada documento; e (A.2) a seriação das unidades espaciais nele referidas (elementos). Se esta seriação implica apenas um subformulário, a que nos referiremos de seguida, já o primeiro bloco contém um conjunto relativamente amplo de campos que é possível agrupar nas seguintes secções:

A.1.1 Identificação: *Nr* – número automático de identificação do documento.

A.1.2 Datação: *Ano*, *Mês*, *Dia* – data cronológica do documento; *AnoF* – ano final do intervalo cronológico do documento, quando aplicável; *Data2* – observações sobre datações críticas, imprecisas, tópicas, etc.; *Séc.*, *50*, *25* – classificação do documento em intervalos de século, meio-século e quarto-de-século.

A.1.3 Informação diplomática: *Fundo*, *Cota* – localização arquivística; *Crítica* – diplomática, não de conteúdo; *Tipologia* – diplomática e não exatamente jurídica; *Tradição* – original ou diversos tipos de cópia; *Produtor* – instituição produtora do documento, quando aplicável.

A.1.4 Publicação: *DocPublicações* – subformulário de que consta a referência às edições modernas do documento e respetivo número de ordem dentro da edição (*Publicação*,

Nr.Publ), bem como a marcação da edição considerada preferencial (*Pref*), nos casos em que haja mais do que uma.

A.1.5 Conteúdo: *Sumário* – resumo breve do conteúdo do documento; *Cit.Doc.* – transcrição das passagens do documento em que são referidas unidades espaciais.

A.1.6 Observações: *Obs.* – notas mais detalhadas sobre informação que consta dos restantes campos ou outros temas que merecem algum tipo de observação.

A.2 Elementos: subformulário que apresenta a lista de todas as unidades espaciais mencionadas no documento, especificando para cada menção: o respetivo número de identificação (*Elemento*); o tipo morfológico e a designação que este documento, em particular, lhe atribui (*Tipo, Designação*); a marcação da unidade como objeto do negócio jurídico consignado pelo documento, quando aplicável (*Obj.*); e a referência ao número de identificação na tabela *Unidades* da unidade espacial a que corresponde cada elemento (*Unidade*). Cada linha deste subformulário identifica um elemento cuja descrição detalhada é registada no próximo módulo, completando os campos aqui presentes com muitos outros.

B) Módulo Elementos

Deste segundo módulo consta um conjunto muito amplo de campos em que se procura registar a informação fornecida pelo documento sobre a morfologia de cada uma das unidades. É possível agrupar estes campos nas seguintes sete secções:

B.1 Identificação²: *Id* – número automático de identificação do elemento; *Tipo* – classificação morfológica atribuída pelo redator do documento à unidade; *Ref. Plural* – marcação das unidades que o documento não individualiza mas refere no plural, como parte integrante de um conjunto de unidades do mesmo tipo; *Designação* – transcrição completa da exata denominação atribuída à unidade no documento. De forma a sistematizar o recurso a estes elementos, definimos seis tipos, que podem coexistir numa mesma designação: *Antroponímico, Proprietário/usufrutuário, Toponímico, Topográfico, Hagionímico, Outros*; *Obj.* – marcação da unidade como objeto do negócio jurídico consignado pelo documento, quando aplicável. Por último, cabem ainda aqui os campos (situados na parte inferior do formulário) de remissão para o número de identificação do diploma em que o elemento é referido (*DocId*) e para a respetiva data (*Ano, AnoF*), bem como para o número de identificação (na tabela *Unidades*) da unidade a que o elemento corresponde. E o campo em que se transcreve detalhadamente as passagens em que o documento alude a esta unidade (*Cit.Doc.*).

¹ O cabeçalho de identificação do elemento (*Id., Tipo, Ref. Plural, Designação*) é complementado, na segunda linha do formulário, pela informação relativa à localização (lugar, freguesia e concelho) e ao tipo e designação normativos da unidade a que cada elemento corresponde, dados provenientes do formulário *Unidades*.

Figura 3: Formulário Elementos

B.2 Fragmentação: *Porções, Frações* – marcação de referências a porções (abstratas) ou frações (devidamente quantificadas) da unidade, normalmente objeto de transação; *Fragm.Obs.* – registo do número de porções e/ou do exato valor das frações referidas, bem como de outros dados relevantes.

B.3 Delimitação: A diversidade de informações que cabem nesta secção obrigou à definição de cinco tipos, que podem coexistir segundo várias combinações possíveis: (i) *Confrontação* – que pode dizer respeito tanto à «Demarcação» linear do perímetro da unidade como à simples indicação de «Elementos confinantes» soltos; ambos os subtipos vão assinalados, com recurso a estas duas expressões, no campo *Del.Obs.*; (ii) *Marcos Mentais* de delimitação; (iii) *Marcos Físicos* de delimitação; (iv) *Dimensões* – referências a dimensões exatas da unidade ou de respetivas porções/frações, que vão registadas no campo *Del.Obs.*; (v) *Outros Del.* – referência a outras formas de delimitação.

B.4 Sistema de localização da propriedade: A relevância da informação espacial veiculada pelo sistema de localização utilizado pelos redatores para situar geograficamente a(s) propriedade(s) objeto de transação justifica a criação de uma secção autónoma em que é possível marcar a referência explícita a um ou mais elementos de localização que compõem esse sistema. Sem esgotar a paleta destes elementos (que podem ser de índole social/territorial ou estritamente física), individualizámos os seguintes, por serem os mais frequentes: *Villa/Top.* (villae ou topónimos); *Montes/Castros*; *Rios*; *Mar*; *Outros EFL* (outros elementos físicos de localização).

B.5 Relações: A preocupação central com a recolha da abundante informação veiculada pela documentação notarial sobre relações espaciais entre unidades levou-nos a dedicar-lhe uma secção autónoma. A distinção essencial entre as partes «livres» e «formulares» do discurso notarial (Sabatini, 1965), com evidentes implicações na qualidade dos dados fornecidos por umas e outras, obrigou à criação de dois subformulários autónomos: B.5.1 *Relações*, em que são arroladas e caracterizadas as relações da unidade em causa com outras unidades dotadas de uma expressão espacial concreta (e por isso também fichadas nos

módulos Elementos e Unidades); e B.5.2 *Relações Formulares/Indefinidas/Menores*, em que são arroladas e caracterizadas as relações da unidade em causa com unidades mencionadas no quadro de enumerações estereotipadas e de outras fórmulas de descrição da propriedade (e que em alguns casos não teriam uma existência concreta, mas um valor de mera plausibilidade), com unidades indefinidas e/ou desprovidas de uma tradução espacial e/ou de uma tipologia concreta, ou ainda com unidades menores (edifícios, águas, etc.) que não valeria a pena fichar autonomamente nos módulos Elementos e Unidades.

B.5.1 *Relações*: subformulário que apresenta a lista das relações, referidas no mesmo documento, entre a unidade em causa e outras unidades, especificando para cada relação: o respetivo número de identificação (*Id.*); o tipo de relação, de acordo com uma tipologia a que nos referiremos de seguida (*Relação*); a identificação do elemento relativo à unidade relacionada (*Unidade2, Tipo, Designação*). Há uma infinidade de tipos de relações possíveis que procurámos sintetizar num conjunto de quatro tipos, a exigirem alguns esclarecimentos: (i) «*integra/integrada*»: aplica-se a todas as unidades que integram outras em si mesmas ou são integradas noutras; (ii) «*confina*»: aplica-se a todas as unidades que confrontam fisicamente umas com as outras; (iii) «*co-integrada*»: aplica-se a todas as unidades que, em virtude da pertença a um mesmo proprietário, da transação conjunta num mesmo ato (mesmo que pertencentes a diferentes titulares) e de outras informações proporcionadas pelo texto, apareçam integradas numa mesma unidade produtiva e/ou residencial; (iv) «*localiza/localizada (...)*»: aplica-se a todas as unidades que são referidas pelos redatores como instrumento («*localiza*») ou objeto («*localizada*») de um qualquer tipo de localização, normalmente assinalada na documentação no quadro do já referido sistema de localização dos bens transacionados ou de meras indicações locativas, um e outras destinados a integrar geográfica e/ou territorialmente as unidades em causa.

B.5.2 *Relações Formulares/Indefinidas/Menores*: subformulário que apresenta a lista das relações mantidas pela unidade em causa com unidades formulares, indefinidas ou menores referidas no mesmo documento, especificando para cada relação: o respetivo número de identificação (*Id.*); o tipo de relação, de acordo com a mesma tipologia definida para o subformulário *Relações* (*Relação*); o tipo morfológico da unidade relacionada (*Tipo*). Porque estas unidades formulares, indefinidas ou menores, não têm uma tradução espacial/material concreta ou não justificam o registo, não foram fichadas como unidades autónomas nos módulos *Elementos* e *Unidades*. Com efeito, o que importa analisar no caso das relações estabelecidas com estas unidades é apenas a morfologia das unidades relacionadas e a tipologia das relações, por forma a definir padrões gerais.

B.6 *Atribuição Social*: O questionário subjacente à BD não contempla apenas a caracterização da morfologia física das unidades espaciais, mas também a da sua morfologia social, para o que é necessário recolher sistematicamente a informação relativa a dois indicadores-chave da «*atribuição social*» destas unidades: a cadeia de transmissão da titularidade sobre essas unidades e o conjunto de imposições que sobre elas recaíam. A estes dois indicadores correspondem dois subformulários: B.6.1 *Transmissões* e B.6.2 *Tributos*.

B.6.1 *Transmissões*: subformulário que apresenta a lista das sucessivas transmissões de que a unidade foi objeto, de acordo com a informação prestada por cada documento, especificando para cada transmissão: o respetivo número de identificação (*Id.*); a identidade do proprietário que aliena a unidade (*Anterior*); a tipologia da transmissão (*Tipo*); a identidade do proprietário que adquire a unidade (*Posterior*); o número de ordem desta transmissão em particular no quadro de uma eventual cadeia de transmissões referida no documento (*Ordem*); a possibilidade de marcar esta transmissão como o negócio jurídico que o ato consigna (*Principal*); e um campo livre de observações (*Obs.*).

B.6.2 Tributos: subformulário que apresenta a lista dos tributos que impendem sobre a unidade, de acordo com a informação prestada por cada documento, especificando para cada tributo: o respetivo número de identificação (*Id.*); a designação exata atribuída pelo documento (*Tributo*); a identidade da autoridade tributária ou simplesmente do indivíduo que determina a imposição e beneficia do tributo (*Autoridade*); a caracterização da situação tributária, de acordo com uma tipologia básica («Exação», «Isenção», «Imunidade») (*Situação*); a indicação do valor do tributo (*Valor*); e um campo livre de observações (*Obs.*).

B.7 Observações: *Obs.* – notas mais detalhadas sobre informação que consta dos restantes campos/secções ou outros temas que merecem algum tipo de esclarecimento particular.

C) Módulo Unidades

Deste terceiro módulo consta a informação geral sobre cada uma das unidades espaciais identificadas, que resulta da agregação de todas as menções documentais a essa unidade (elementos). Os vários campos que compõem este módulo dividem-se em três grandes secções:

C.1 Identificação: *Id* – número automático de identificação da unidade; *TipoNorm* – classificação morfológica normativa, que resulta da ponderação das classificações, nem sempre coincidentes, atribuídas à unidade nos diversos documentos em que é mencionada; *Desig-Norm* – denominação normativa, que reúne os vários elementos de designação da unidade referidos nos diversos documentos que a mencionam.

O problema central que se coloca nesta secção é o da classificação morfológica, uma vez que o questionário subjacente à BD coloca no centro da análise os diversos tipos de unidades de organização do espaço definidos pelos redatores dos documentos, com recurso a um léxico classificatório que se constrói na interseção entre os planos material e jurídico e que responde a esquemas sociais de organização e a esquemas mentais de representação do espaço. A profusão de termos integrados neste léxico espacial obrigou-nos, desde logo, a agregar alguns desses termos em tipos compósitos (que reúnem duas palavras com sentidos próximos mas não coincidentes), sintéticos (que congregam unidades semelhantes do ponto de vista morfológico mas designadas por palavras diferentes) e residuais («Outros»). E ditou sobretudo a necessidade de criar uma tipologia que pudesse, de alguma forma, sistematizar essa profusão de termos/tipos, agrupando-os em categorias definidas por critérios que passam pela morfologia das unidades, em primeiro lugar, mas também pelo enquadramento discursivo mais frequente desses termos (partes do discurso diplomático e tipologias documentais em que ocorrem) ou mesmo, no caso das unidades de paisagem, os diversos setores a que pertencem. Note-se, todavia, que se em alguns casos foi necessário criar categorias de primeira, segunda e terceira ordem para arrumar conjuntos de tipos muito numerosos e morfológicamente variados, noutros não se ultrapassou a categoria de primeiro nível.

FIGURA 4 – Formulário Unidades

C.1

C.2

Id	Tipo	Designação	Doc.	Ano
3193	Villa	uilla lalini	338	959
3293	Villa	uilla lalini	341	960
3577	Villa	Villa de lalini	356	1014
3842	Villa	villa lalini integra cum suo mandamento	381	1059
4539	Villa	[villa] Larim cum abiuntionibus suis	394	950
*	(Novo)			

C.3

C.4

OBS - LOCALIZAÇÃO: Esta villa é unanimemente situada no l. Larim da f. Soutelo: COSTA, 2000: 192; RAMOS, 1991: 112. Anexo ao mapa n.º 4; AMARAL, 2007: 615, que esclarece: «L. da f. de S. Miguel de Soutelo, primitivamente chamada de Larim, c. de Vila Verde». Note-se contudo que o topónimo "Larim" também se aplicou à f. Lourera (Sta. Eulália) «antigamente Larim ou Vila Verde», segundo COSTA, 2000: 626. Poderá deduzi-se daqui que a villa primitiva ocuparia um perímetro hoje repartido entre as ambas as fs. (vizinhas)? Esta villa distingue-se ia ainda da villa Soutelo (v. Obs. à un. 1853). - ELEMENTOS: el. 4539; FERNANDES, 1970-1972, II (1971): 53 recusa a identificação deste el. com a f. Soutelo, c. Vila Verde: «Larim, não o do Cávado, mas em Deza [Galiza]». No entanto, tanto COSTA como AMARAL arrolam explicitamente o el. 4539 como a primeira menção documental à villa de Larim. É verdade que esta villa veio a pertencer a D.ª Chãmoa Rodrigues, sobrinha de D. Mumadona, que a doou a Gumarões em 960 (doc. 341), e que parecia ter no villa de Larim o centro de um domínio mais vasto concentrado no curso final do Homem, junto à margem do Cávado (v. Notas a documentos específicos; §9.3). Vê-se assim com mais dificuldade como é que esta villa teria cabido em partilhas a D.ª Ónega, filha de Mumadona, apenas 10 anos antes de D.ª Chãmoa a doar ao mosteiro. É este deve ter sido o argumento (não explicitado) que levou A. Fernandes a propor antes a identificação do topónimo com um l. galego. No entanto, não há nenhuma indicação no doc. 354 que permita situar esta villa na Galiza. E convém mesmo notar, embora o argumento seja débil, que os ls. arrolados imediatamente antes e depois de "Larim", entre os bens que couberam a D.ª Ónega, estão todos localizados no território portuguêsense. - FUNÇÕES/RESIDENCIAIS/PRODUTIVAS: el. 3577: «Villa de lalini et souteio cum adiuntionibus suis et omnes habitantes in eas stue uillas quomodo ei incommunitas ab integro sicut in scripturis colligati sunt» (doc. 356).

C.2 Localização: *Id Lugar* – número automático de identificação do atual lugar em que a unidade se situa; *Concelho, Freguesia, Lugar* – indicação dos atuais concelho, freguesia e lugar em que a unidade se situa. A informação que consta destes campos provém da Base Geográfica de Referência da Informação (BGRI²), produzida pelo INE, de que foi extraída para a BD a informação relativa aos atuais distritos de Braga, Bragança, Viana do Castelo e Vila Real. Embora estas listas de concelhos e freguesias sejam à partida exaustivas, já a de lugares tem de admitir a adição de atuais lugares em falta e sobretudo da multiplicidade de antigos lugares mencionados na documentação que ou desapareceram ou nos foi impossível identificar.

O grande problema que se coloca nesta secção é o da identificação toponímica das unidades. Naturalmente, há que admitir a possibilidade de erro puro e simples, decorrente da falta de elementos informativos que sustentem identificações com total acerto, ou mesmo da pura deslocação dos núcleos de habitat entre o período em estudo e a atualidade. No entanto, importa sobretudo chamar a atenção para a possibilidade de erro, ditada pelas diferenças estruturais entre a base territorial e o sistema de localização utilizados pelos redatores dos documentos para nomear e localizar as unidades espaciais e a malha administrativa atual que informa a nossa perceção espacial e a que está obrigada qualquer tentativa de criação de uma malha territorial homogénea capaz de suportar a georreferenciação destas unidades, imprescindível desde logo à sua representação cartográfica (Bourin & Zadora-Rio, 2007, pp.45-46).

C.3 Elementos: subformulário que apresenta a lista de todas as menções documentais à unidade em causa (*elementos*), especificando para cada menção: o respetivo número de

² A BGRI (2001) esteve na raiz da implementação do suporte digital para a cartografia censitária: http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos_base_cartogr (cons. 2012-07-14).

identificação na tabela *Elementos (Id)*; o tipo morfológico e a designação que o respetivo documento, em particular, lhe atribui (*Tipo, Designação*); o número de identificação do documento em que cada elemento é mencionado (*Doc.*) e o respetivo ano, essencial para a datação dos vários elementos relativos a uma mesma unidade (*Ano*).

Como já ficou dito, a identificação das diversas menções documentais a uma mesma unidade está longe de ser uma operação automática. Pelo contrário, a homonímia que caracteriza a toponímia do período e da documentação estudados, bem como a escassez da informação relativa a muitas das unidades referidas (sobretudo quando não são objeto dos atos jurídicos consignados pelos documentos), dificultam grandemente esta operação. No essencial, ela assenta em três tipos de dados: (i) a designação, que em alguns casos se limita a um topónimo, antropónimo determinativo, hagiónimo, etc., que pode não bastar para distinguir entre duas unidades com a mesma designação e/ou pertencentes a um mesmo proprietário/usufrutuário e/ou situadas na mesma área; (ii) a informação relativa à titularidade das unidades, que pode ir da mera identidade dos proprietários/usufrutuários (atuais e/ou anteriores) que as transacionam até ao estabelecimento de verdadeiras cadeias de transmissão dessa titularidade ou mesmo a referência a porções/frações que denunciem a repartição (hereditária ou não) de uma mesma unidade; e (iii) todos os dados fornecidos pelos redatores dos documentos para localizar as unidades referidas, desde os elementos toponímicos (ou topográficos) de designação, até à menção a macro e micro-elementos de localização (com destaque para os que integram o já referido sistema estruturado: terras/territórios, villae/topónimos, montes/castros, rios, mar, etc.), ou mesmo a referência a unidades confinantes e outra informação que consta da delimitação das unidades.

C.4 Observações: *Obs.:* notas mais detalhadas sobre informação que consta dos restantes campos ou outros temas que merecem algum tipo de observação.

4. EXPLORAÇÃO DA BASE DE DADOS

As muitas possibilidades de tratamento abertas pela recolha sistematizada da informação numa BD intencionalmente relacional estão na raiz de um conjunto alargado de potencialidades.

4.1 RESULTADOS OBTIDOS

Respondendo aos objetivos inicialmente enunciados, a BD foi dotada de um conjunto de mecanismos analíticos que procuram viabilizá-los. Atentemos aqui em cinco tipos de resultados:

- (i) Elaborou-se um conjunto de consultas de verificação da qualidade dos dados que permitiu identificar e corrigir erros de recolha, como por exemplo, a falta de reciprocidade no estabelecimento de relações entre elementos.
- (ii) Para assegurar a exaustividade da análise, a BD prevê a produção de listagens exaustivas dos registos efetuados, segundo vários critérios. A Figura 5 mostra a listagem das unidades, incluindo a informação de lugar em que se situam.

FIGURA 5 – Listagem de unidades com localização

Unidades.Id	Tipo	DesignNorm	LugarFre	Lugar	Freguesia	Concelho
1174	Ager, Campo	Agra	10169 s.l.	Lage	Vila Verde	Vila Verde
1302	Ager, Campo	Agra	10174 s.l.	Moure	Vila Verde	Vila Verde
409	Ager, Campo	Agra	17078 Speranzo (Tulio?) (a	Passos (S Julião)	Braga	Braga
1325	Ager, Campo	Agra de Arca	4192 Arca	Turiz	Vila Verde	Vila Verde
1257	Ager, Campo	Agra de Libão	10169 s.l.	Lage	Vila Verde	Vila Verde
1651	Ager, Campo	Agra de S. Cristóvão	9818 s.l.	Esporões	Braga	Braga
1451	Ager, Campo	Agra de sob S. Salvador	3823 Parada	Barbudo	Vila Verde	Vila Verde
1211	Ager, Campo	Agra de Zam	10174 s.l.	Moure	Vila Verde	Vila Verde
178	Ager, Campo	Agra dita Várzea	17332 Várzea (a.l.?)	Este (S Mamede)	Braga	Braga
357	Ager, Campo	Agra Fonte	9824 s.l.	Gualtar	Braga	Braga
416	Ager, Campo	Agra Molino	13665 s.l.	s.f.	Braga	Braga
2549	Ager, Campo	Agraza	12162 s.l.	Vilarinho	Santo Tirso	Santo Tirso
2445	Ager, Campo	Agrela	12162 s.l.	Vilarinho	Santo Tirso	Santo Tirso
3229	Ager, Campo	Agrela	10113 s.l.	Gavião	Vila Nova de F	Vila Nova de F
939	Ager, Campo	Agrelo	9826 s.l.	Lamações	Braga	Braga

(iii) A Figura 6 mostra um conjunto de indicadores estatísticos: para cada tipo normativo calcula-se o número de unidades que verifica cada uma das variáveis de classificação morfológica dos respetivos elementos.

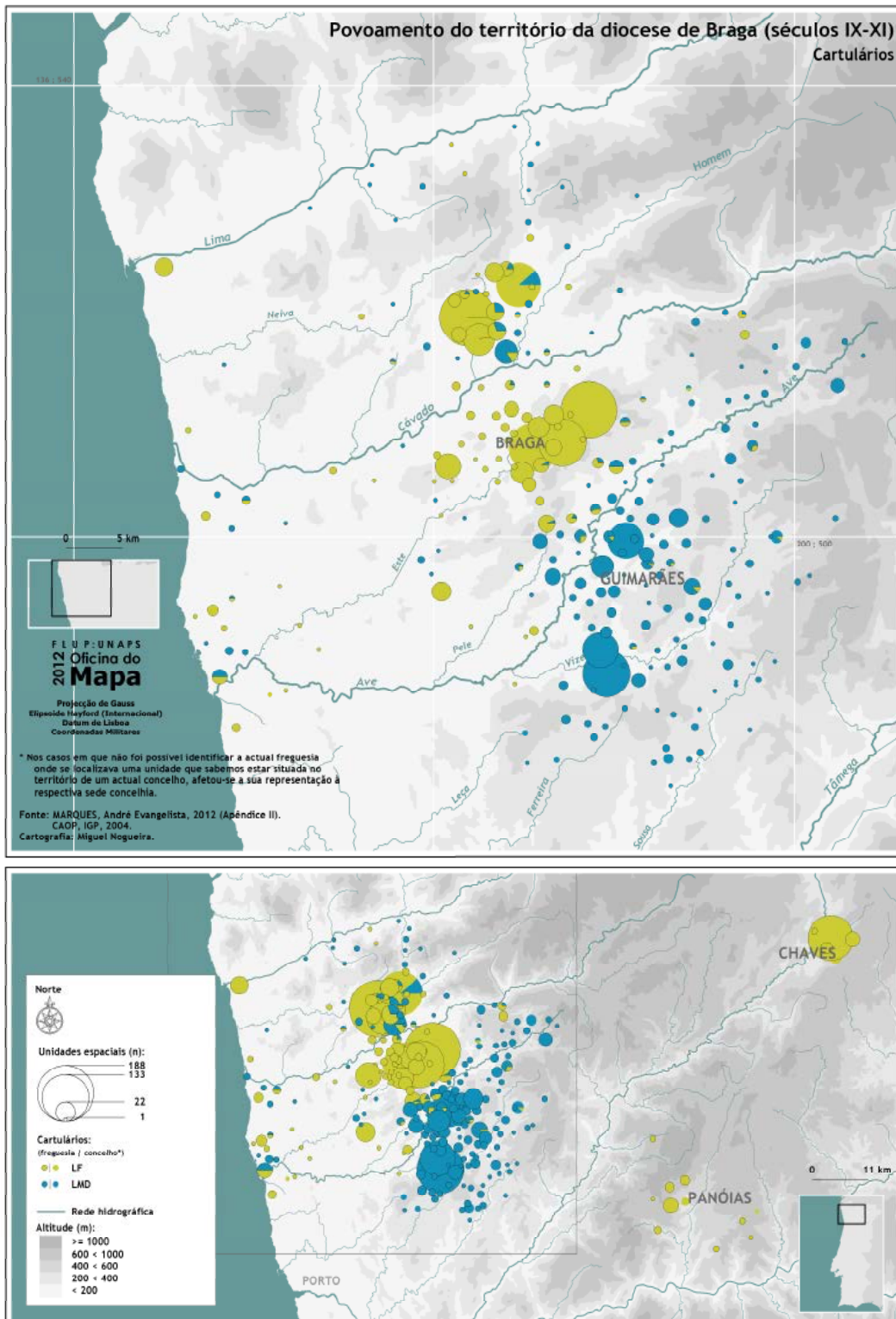
FIGURA 6 – Frequência de atributos por tipo de unidade

Tipo	Categoria	Categoria2	Categoria3	Ref Plural	Antroponí	Proprietáric	Toponím	Topográf	Hagioním	Outros EID
Villa	UOSE	UOSE/Pov	UOSE/Pov	4	10	24	283	8	19	5
Castrum	UP	UP/F	UP/F	0	0	2	11	0	2	4
Casal	UOSE	UOSE/Exp	UOSE/Exp	4	54	89	6	21	0	0
Casa	UP	UP/R	UP/R	2	22	6	4	1	1	2
Domus	UP	UP/R	UP/R	0	7	3	0	0	0	0
Quintana/quinta	UOSE	UOSE/Exp	UOSE/Exp	0	0	4	1	0	0	1
Hereditas	FP	FP	FP	2	11	66	2	7	0	0
Vicus	UOSE	UOSE/Pov	UOSE/Pov	0	0	0	2	0	0	0
Territorium (diocesi)	UASE	UASE	UASE	0	0	0	1	0	0	1
Terra, Territorium	UASE	UASE	UASE	0	0	1	14	1	1	0
Urbs, Civitas	UASE	UASE	UASE	0	0	0	4	0	0	0
Ecclesia	UE	UE/E	UE/E	0	0	2	49	5	102	2
Mosteiros	UE	UE/M	UE/M	0	0	2	21	3	17	1
Templos (outras de)	UE	UE/T	UE/T(o/d)	0	0	0	3	0	4	0
Archidiaconatus	UASE	UASE	UASE	0	0	1	0	1	0	0

(iv) Para assegurar a rigorosa georreferenciação de toda a informação espacial, com vista à sua integração num SIG, as listas de lugares e freguesias que servem de base à identificação toponímica das unidades foram dotadas das respetivas coordenadas geográficas. A partir daí produziram-se consultas em formato adequado à produção automática de cartografia de que se apresenta na Figura 7 um exemplo: a distribuição do total de unidades espaciais identificadas por freguesia em cada um dos cartulários analisados.

(v) A BD constitui ela própria um resultado do trabalho efetuado, no sentido em que é disponibilizada para estudos posteriores num repositório científico de acesso aberto, fixando um corpus de unidades pesquisável em função de coordenadas espaciais, cronológicas ou morfológicas, devidamente documentado e justificado.

FIGURA 7 – Distribuição das unidades espaciais documentadas em cada um dos cartulários analisados



4.2 POTENCIALIDADES ANALÍTICAS

Análise quantitativa

As limitações informativas e os inúmeros problemas de crítica inerentes à documentação altimedieval, bem como as pronunciadas descontinuidades cronológicas e espaciais dos corpora disponíveis para este período, dificultam naturalmente uma abordagem estatística ou «serial» deste tipo de fontes. Coloca-se, portanto, um problema de dimensão e representatividade das amostras. Todavia, sem prejuízo destas limitações das fontes medievais para o estabelecimento de séries de dados passíveis de uma análise verdadeiramente estatística, é inegável o interesse heurístico da quantificação, desde logo no campo dos estudos semânticos (GUERREAU, 2001, p.304; ARNOLD, 2008, pp.55, 65 e ss.). Uma vez analisadas e devidamente ponderadas as limitações informativas, os problemas de crítica e as descontinuidades do corpus documental, com vista à definição da representatividade da amostra, é possível recorrer à análise quantitativa dos dados por forma a inferir distribuições, variações e padrões que, sendo relativos, não deixam de ter alguma relevância intrínseca, quando não são mesmo capazes de indicar tendências mais gerais, que ultrapassam os limites da amostra. Aliás, a crítica das fontes assume-se como uma condição imprescindível e fundadora da análise quantitativa, que por sua vez permite avançar no esclarecimento dos significados subjacentes às palavras.

No entanto, há no tipo de fontes analisadas um conjunto bem identificável de dados que, em virtude da frequência/repetição com que a documentação no-los apresenta, são passíveis de uma análise quantitativa. O método que aqui propomos procura criar as condições para um aproveitamento quantitativo dos dados recolhidos em documentos que são tradicionalmente vistos como sendo hostis a qualquer tipo de quantificação. A informação fornecida pela documentação diplomática altimedieval é certamente escassa e fragmentária. Todavia, o volume de dados que resulta de levantamentos exaustivos pode atingir uma dimensão e qualidade bem superiores àquilo que normalmente se esperaria. A realidade espacial é um dos domínios em que isso se verifica (NOËL, 2010, pp.40-41). O que se deduz claramente do corpus documental que está na base deste trabalho: num total de 366 documentos analisados, foi possível identificar 3073 unidades espaciais, a que correspondem 4937 menções documentais (*elementos*), entre as quais foi possível estabelecer um total de 11516 relações espaciais. Seria impossível manejar este corpo de dados sem recurso à análise quantitativa, independentemente da utilização ou não de técnicas estatísticas propriamente ditas, com destaque para a análise multidimensional das diversas variáveis contempladas no questionário que estrutura a nossa metodologia (GUERREAU, 2001, pp.179-81).

Note-se, contudo, que a análise quantitativa não tem apenas a virtude de tornar manejável um corpo tão vasto de informação e de permitir identificar as respetivas distribuições e variações. Quando agregada, essa informação revela padrões e tendências mais ou menos gerais ou generalizáveis que, em si mesmas, possibilitam leituras interpretativas totalmente vedadas à análise qualitativa, por natureza obrigada à especificidade dos dados individualmente considerados.

Integração dos dados num SIG: cartografia e análise espacial

Enquanto «proposta de análise espacial da documentação altimedieval», o método aqui apresentado não poderia deixar de estar orientado para a integração entre a BD construída para a recolha e tratamento dos dados provenientes das fontes escritas e uma base geográfica. A integração entre os dois tipos de BD implica, como ponte entre uma e outra, a cuidadosa georreferenciação de toda a informação, que deve seguir uma escala tão fina quanto

possível (NOGUEIRA, 2010). Só assim será possível avançar na análise relacional entre a informação histórica e todas as variáveis geográficas (físicas como humanas) que com ela mantenham um qualquer tipo de associação significativa.

Naturalmente, e apesar das muitas possibilidades de análise espacial abertas por qualquer SIG, é na cartografia dos dados produzidos no cruzamento entre ambas as bases (histórica e geográfica) que reside o instrumento primeiro de análise. Impõe-se que os estudos construídos sobre este tipo de fontes ultrapassem definitivamente o recurso à cartografia como mero meio de demonstração de resultados, para a utilizarem como instrumento heurístico. Esta é uma das reivindicações essenciais da presente metodologia, mesmo reconhecendo as limitações que a construção (mais do que tradução) cartográfica dos dados comporta.

Com efeito, insistir na relevância que deve ser atribuída à cartografia implica desenvolver um conjunto de problemas inerentes à elaboração cartográfica deste tipo de informações. Em primeiro lugar, estão os limites da informação documental disponível. Em segundo lugar, vêm as dificuldades de identificação toponímica, em virtude da simples falta de informação, da incompatibilidade estrutural entre as malhas territoriais subjacentes aos documentos e a atual malha administrativa, das alterações toponímicas/topográficas verificadas entre umas e outra, etc. Finalmente, coloca-se o problema determinante da escolha da unidade espacial de cartografia (fator essencial na hora de definir a escala de representação), da qual dependem em boa medida o alcance e as potencialidades do mapa, logo da própria análise que nele assenta (NOGUEIRA, 2010).

O cuidado posto na identificação toponímica tão rigorosa quanto possível de cada unidade e a preocupação de associar a lista de lugares utilizada na BD às respetivas coordenadas geográficas criam todas as condições necessárias à produção de cartografia em diferentes escalas.

5. CONCLUSÕES

A «densidade» da informação espacial recolhida segundo a metodologia aqui proposta é suficiente para garantir a possibilidade de as análises monográficas dos diversos tipos de unidades espaciais se conjugarem numa visão global capaz de superar a fragmentação a que uma metodologia tão exaustiva quanto possível de recolha de dados obrigou. Mas que, como contrapartida, devolve um manancial de informação sobre cada unidade que não é só abundante como tem a virtude de estar rigorosamente referenciado, tanto do ponto de vista cronológico (uma vez que todo e qualquer dado está indexado ao exato documento que o menciona) como geográfico (dentro, obviamente, das possibilidades de referência oferecidas pelos documentos). Será assim possível, a partir do momento em que esse manancial de informação tenha sido minimamente tratado e analisado numa perspetiva monográfica, avançar para o estudo global da organização da paisagem e do povoamento no quadro regional definido.

Uma análise assim conduzida oferece-nos a possibilidade de indexar um conjunto alargado de informações de vária natureza e rigorosamente datadas (*elementos*) a uma realidade espacialmente circunscrita (a unidade espacial) e, acima de tudo, descrita morfologicamente por meio de um sistema de classificação que, com todas as suas ambiguidades e opacidades para o historiador, e apesar da oscilação entre os planos material e jurídico, tem a imensa vantagem de ser coevo da realidade espacial que procura representar (mais do que propriamente descrever). Ora, é precisamente na especificidade da informação assim produzida que reside o potencial heurístico e interdisciplinar da metodologia aqui apresen-

tada. Superando (sem as anular) dicotomias como as que opõem os planos da representação vs. materialidade do espaço, ou a análise qualitativa de realidades concretas vs. análise quantitativa de agregados abstratos, o carácter individual das unidades espaciais erigidas em unidades de análise garante a possibilidade de ancorar a abstração das palavras no terreno e de contrastar as tendências e ordens de grandeza estatísticas com o caso singular. Num movimento de permanente vai-e-vem, estes vários tipos de análises contrafortam-se uns aos outros.

Para além da publicação destes resultados da investigação sob a forma de tese de doutoramento, a BD desenvolvida constitui uma publicação científica em si mesma, dado o conhecimento especializado necessário para organizar, recolher, filtrar, classificar e analisar os dados que a integram. A identificação das unidades espaciais a partir das fontes documentais é disso um exemplo imediato. A importância desta publicação num repositório de dados científicos é dupla: dá oportunidade à verificação dos resultados por parte da comunidade e permite o desenvolvimento de mais investigação sobre os mesmos dados primários. Esta perspectiva está alinhada com o que se tem designado por e-Science, investigação realizada sobre fontes de dados confiáveis presentes na Internet.

Este trabalho representa mais uma instância da metáfora do computador como telescópio da complexidade. A abordagem ao registo dos dados recolhidos numa BD relacional e o método utilizado para a construir provaram ser um meio de registar relações extensas e complexas e de explorar diferentes linhas de raciocínio em História como noutros domínios do conhecimento.

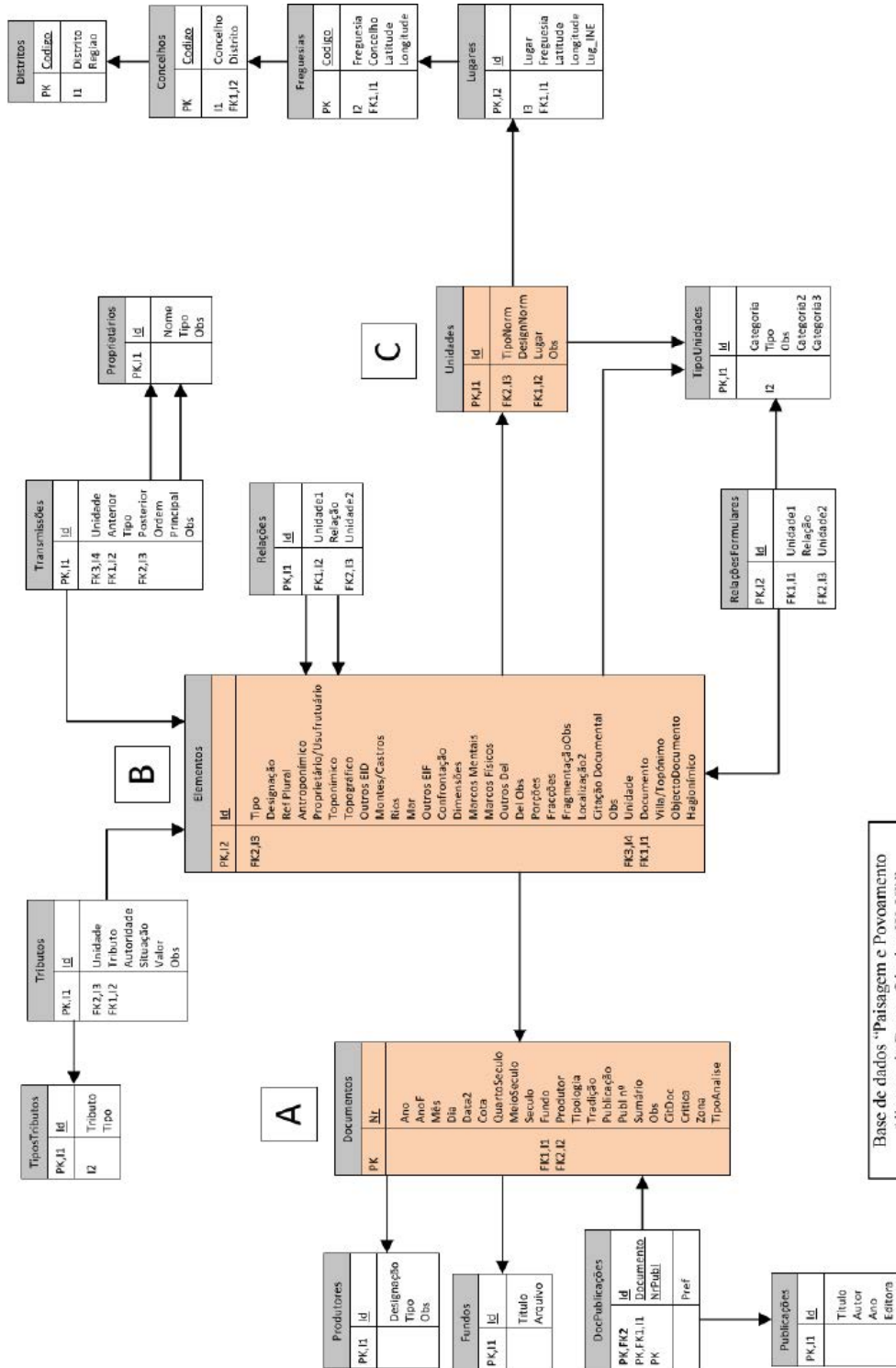
REFERÊNCIAS

- Arnold, John H. (2008). *What is Medieval History?* Cambridge: Polity Press.
- Booch, Grady, Rumbaugh, James & Jacobson, Ivar (1999). *The unified modeling language: user guide*. Addison Wesley, Reading, MA; ISBN 0-201-57168-4; 482 pp.
- Bourin, Monique; Zadora-Rio, Élisabeth (2007). *Pratiques de l'espace: les apports compares des données textuelles et archéologiques*. In *Construction de l'Espace au Moyen Âge: pratiques et représentations. XXXVIIe Congrès de la SHMES, Mulhouse, 2-4 juin 2006* (pp. 39-55). Paris: Publications de la Sorbonne.
- García de Cortázar, José Ángel (1988). *Organización social del espacio: propuestas de reflexión y análisis histórico de sus unidades en la España medieval*. *Studia Histórica-Historia Medieval*, VI, 195-236. Reed. in Idem (2004) – *Sociedad y organización del espacio en la España medieval* (pp. 59-136). Granada: Editorial Universidad de Granada.
- García de Cortázar, José Ángel (1998). *Sociedad y organización social del espacio castellano en los siglos VII al XII. Una revisión historiográfica*. In Hidalgo, M.^a José; Pérez, Dionisio; Gervás, J. R. (eds.). «Romanización» y «Reconquista» en la Península Ibérica: *Nuevas Perspectivas* (pp. 317-337). Salamanca: Ediciones Universidad de Salamanca.
- García de Cortázar, José Ángel (1999). *Organización del espacio, organización del poder entre el Cantábrico y el Duero en los siglos VIII a XIII*. In Idem (ed.), *Del Cantábrico al Duero: Trece estudios sobre organización social del espacio en los siglos VIII a XIII* (pp. 15-48). [Santander]: Universidad de Cantabria; Parlamento de Cantabria. Reed. in Idem (2004) – *Sociedad y organización del espacio en la España medieval* (pp. 140-184). Granada: Editorial Universidad de Granada.
- García de Cortázar, José Ángel (2004). *Sociedad y organización del espacio en la España medieval*. Granada: Editorial Universidad de Granada.
- Guerreau, Alain (2001). *L'avenir d'un passé incertain. Quelle histoire du Moyen Age au XXI^e siècle?* Paris: Éditions du Seuil.

- Noël, René (2010), À la recherche du village médiéval. Hier et aujourd'hui. In Yante, J.-M.; Bultot-Verleysen, A.-M. (eds.). *Autour du «village». Etablissements humains, finages et communautés rurales entre Seine et Rhin (IVe-XIIIe siècles)*. Actes du colloque international de Louvain-la-Neuve, 16-17 mai 2003 (pp. 3-75). Louvain-la-Neuve: Institut d'études médiévales de l'Université Catholique de Louvain.
- Nogueira, Miguel (2010) – Percurso metodológico para a implementação de um SIG em arqueologia mineira: breves reflexões. In Martins, C. M. B (coord.). *Mineração e povoamento na Antiguidade no Alto Trás-os-Montes Ocidental* (pp. 179-187). Porto: CITCEM.
- Peña Bocos, Esther (1995). *La atribución Social del Espacio en la Castilla Altomedieval. Una Nueva Aproximación al Feudalismo Peninsular*. Santander: Universidad de Cantabria, Asamblea Regional de Cantabria.
- Sabatini, Francesco (1965). Esigenze di realismo e dislocazione morfologica in testi preromani. *Rivista di Cultura Classica e Medievale*, 7, 972-998.

ANEXO A

DIAGRAMA RELACIONAL DA BASE DE DADOS «PAISAGEM E POVOAMENTO (DIOCESE DE BRAGA, SÉCULOS IX-XI)»



Base de dados «Paisagem e Povoamento (diocese de Braga, Séculos IX-XI)»
Modelo de dados