

*O Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural  
Cabeceiras de Basto, 23 a 25 de Outubro de 2003*

## **DESENVOLVIMENTO SUSTENTADO - PATRIMÓNIO ARQUEOLÓGICO E TECNOLOGIAS DE INFORMAÇÃO**

### **Natália Botica**

Unidade de Arqueologia da Universidade do Minho, Braga, Portugal  
nb@uaum.uminho.pt

### **Francisco Sande Lemos**

Unidade de Arqueologia da Universidade do Minho, Braga, Portugal  
lemos@uaum.uminho.pt

### **Maribel Yasmina Santos**

Departamento de Sistemas de Informação da Universidade do Minho, Guimarães, Portugal  
maribel@dsi.uminho.pt

## **DESENVOLVIMENTO SUSTENTADO - PATRIMÓNIO ARQUEOLÓGICO E TECNOLOGIAS DE INFORMAÇÃO**

### **RESUMO**

As sociedades estão cada vez mais motivadas para preservarem os testemunhos do passado.

Essa atitude, consubstanciada em políticas, leis, instituições e actividades tem gerado várias iniciativas orientadas para a preservação e valorização do património, como forma de salvar os testemunhos materiais do passado.

As Tecnologias de Informação são relevantes na concretização desta tarefa. Apoiam a gestão dos dados associados ao património arqueológico e garantem o suporte a trabalhos de pesquisa e de divulgação dos valores patrimoniais.

Estas tecnologias permitem ainda desenvolver modelos para elaboração de cartas arqueológicas, delimitando áreas de risco para salvaguarda de sítios visíveis. O património oculto, que surge muitas vezes onde menos se espera, pode também ser detectado com base nas ferramentas de Descoberta de Conhecimento em Base de Dados e cartografado em função dos indicadores dos contextos e características de sítios já referenciados.

**PALAVRAS CHAVE:** património arqueológico, tecnologias de informação, descoberta de conhecimento em bases de dados.

### **1. INTRODUÇÃO**

Os ecossistemas humanizados nos quais se inserem um universo de numerosas formas de património sofreram e continuam a sofrer sucessivas agressões. O tempo, aliado ao crescimento demográfico e outras alterações da paisagem decorrentes das actividades

industriais, agrícolas e florestais, constituem linhas de desenvolvimento que têm obliterado muito património.

Despontam, no entanto, algumas louváveis iniciativas nas áreas do turismo cultural, ecológico ou histórico, orientadas para recreação de espaços e vivências do passado. É uma nova economia, cujo principal recurso são as paisagens culturais, que deverá manter vivas, evitando a sua degradação e desaparecimento.

O desenvolvimento destas actividades poderá contribuir de forma decisiva para que o património seja estudado, preservado e valorizado. Também as Tecnologias da Informação têm assumido uma importância crescente em todo o processo de registo e estudo dos valores patrimoniais. Começaram por proporcionar o uso de bases de dados onde se armazenava e geria a informação sobre os sítios arqueológicos. Com base nessa informação apoiaram de forma decisiva os projectos de investigação, o desenvolvimento de Sistemas de Informação Geográfica, a reconstituição virtual de património e a criação de conhecimento arqueológico.

Este artigo encontra-se organizado da seguinte forma: na secção 2 apresenta-se um sistema de inventariação de Património, onde se procede ao registo e gestão da informação dos sítios arqueológicos, aplicada ao Concelho de Fafe. Na secção 3 utiliza-se uma base de dados de arqueossítios para desenvolver um modelo de apoio à prospecção de sítios arqueológicos. O exemplo apresentado é o sistema de apoio à prospecção arqueológica para a região de Trás-os-Montes, onde se utiliza a Descoberta de Conhecimento em Bases de Dados.

## **2. GESTÃO DA INFORMAÇÃO DE PATRIMÓNIO ARQUEOLÓGICO**

O sistema de gestão de informação do património arqueológico a seguir apresentado teve por base um trabalho realizado na Unidade de Arqueologia da Universidade do Minho (UAUM) a fim de rever a Carta do Património Arqueológico do PDM (Plano Director Municipal).

As Cartas Arqueológicas, sendo um inventário o mais exaustivo possível dos sítios conhecidos, constituem um instrumento prévio fundamental para se definir uma política patrimonial e proceder à adequada gestão dos bens arqueológicos.

A abordagem feita pela equipa pluridisciplinar da UAUM para proceder ao inventário do património arqueológico passou pela realização de diversas tarefas, ressaltando-se:

- o estudo bibliográfico e documental de todos os trabalhos anteriores, com referências ao Concelho, incluindo o inventário arqueológico realizado em 1983;
- o estudo da cartografia da região;
- a definição de formulários a utilizar em campo;
- os trabalhos de prospecção no terreno;
- a análise e tratamento da informação recolhida;
- a proposta de zonamentos arqueológicos;
- a elaboração de memória científica.

Durante todo o processo foram utilizadas Tecnologias da Informação. Pretendeu-se criar um sistema capaz de reunir, guardar, processar e facultar dados relevantes para a concretização dos objectivos definidos, mas que fosse também fonte de dados para outros trabalhos, nomeadamente a criação de uma página na Internet, a produção de conteúdos multimédia, a elaboração de folhetos de divulgação de património ou de roteiros culturais.

A aplicação de Gestão da Informação do Património Arqueológico de Fafe (GIPAF), assenta numa Base de Dados Relacional Oracle, a funcionar numa arquitectura cliente-servidor e possui três componentes base:

- Base de Dados;
- Base Cartográfica;
- Banco de Imagens.

## 2.1. BASE DE DADOS

O inventário dos sítios do Concelho inclui a caracterização de 265 arqueossítios, dos quais cerca de 100 surgiram como resultado deste trabalho, não constando da anterior carta arqueológica (1983). Muitos destes locais são passíveis de classificação e integráveis em itinerários culturais e turísticos.

O inventário do património foi precedido de um levantamento das necessidades de informação, desenharam-se formulários para levantamento de dados (figura 1) e foi criada uma base de dados, para registo dos elementos recolhidos durante as deslocações aos sítios e nas pesquisas bibliográficas.

The image displays two forms for archaeological site registration. The left form is titled 'UAUM' and includes sections for 'Localização' (Location) with fields for 'Topónimo', 'Lugar', 'Freguesia', 'Concelho', and 'Distrito'; 'Coordenadas' (Coordinates) with 'Cartografia' and 'GAUSS X', 'GAUSS Y', and 'Altitude' fields; 'Acessos' (Accesses); and 'Zonas Envolventes' (Surrounding Areas). The right form is titled 'UAUM' and includes sections for 'Registo de material' (Material Record) with 'Número de peças', 'Assunto', 'Assunto secundário', 'Material empilhado', and 'Assunto de estudo'; 'Conteúdo do sítio' (Site Content) with 'Tipo de trabalho realizado' and 'Material encontrado'; and 'Sitio' (Site) with 'Designação', 'Contexto', 'Descrição', 'Estado de conservação', and 'Valorização'. Two orange arrows indicate the flow of information between the forms.

Figura 1 – Formulários de registo de património arqueológico

Como interface do utilizador com a Base de Dados desenvolveu-se uma aplicação para gestão do material informativo. A aplicação denominada GIPAF, da qual se apresentam alguns ecrãs na figura 2, permite gerir os dados, estruturados de acordo com as várias componentes definidas: Identificação do Bem, Localização, Registo de Estruturas e Materiais, Estado de Conservação, Situação e Valorização Arqueológica e Patrimonial.

*O Congresso Internacional de Investigação e Desenvolvimento Sócio-cultural  
Cabeceiras de Basto, 23 a 25 de Outubro de 2003*

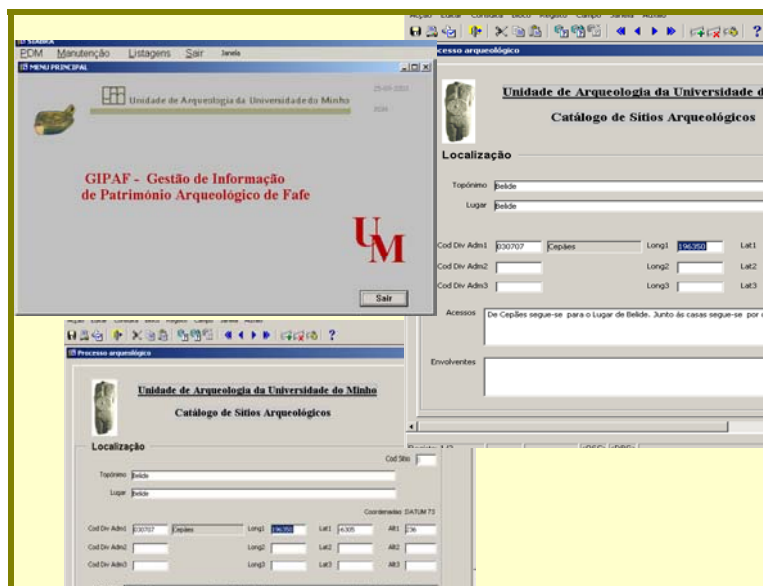


Figura 2 – Ecrãs da aplicação de gestão da informação do GIPAF

O valor e utilidade da Informação não está apenas nos dados mas na utilização que lhes é dada. Tendo em mente esta premissa, definiram-se algumas “queries” para interrogação da Base de Dados e pré desenharam-se alguns relatórios, estruturados por tipo de sítios e por freguesias. No entanto, há uma grande diversidade de reports possíveis e que podem ser elaborados de acordo com as necessidades dos utilizadores.

Sítio	Valorização Patrimonial	Valorização Arqueológica	Estado Conservação
Mamos 01 do Marco	Tem acessos fáceis. Poderá ser sinalizada e enquadrado num roteiro turístico.	Este Mamos pertence a uma necrópole megalítica.	Médio
Mamos 02 do Marco	Os acessos são fáceis. Poderá ser sinalizada e enquadrado num roteiro turístico.	Este mamos pertence a uma necrópole megalítica.	Bom
Mamos 01 Campo Salgueiro	Acessos fáceis. Poderá ser sinalizada e enquadrado num roteiro turístico.	Este mamos pertence à mesma necrópole das Mamos I e II do Marco.	Bom
Mamos 01 Monte da Cumieira	Acesso Fácil. A mamos localiza-se junto à estrada.	Poderá ter interesse científico para futura relação com outros monumentos do tipo, visto estar inserido num planalto onde já estão referenciados outros monumentos.	Bom
	Acessos fáceis. Inserido num núcleo megalítico que seria importante sinalizar e valorizar.	Pertence à mesma necrópole megalítica das Mamos do Lameiras.	Bom
	Acessos fáceis. A necrópole megalítica poderá ser sinalizada e enquadrado num roteiro turístico.	Pertence à mesma necrópole megalítica das Mamos do Lameiras.	Bom
	Acessos fáceis.	A mamos foi destruída pelo espoliação de sobras. Não se recolheu nenhum espólio.	Mau

Figura 3 – Listagens do GIPAF

## 2.2. BASE CARTOGRÁFICA

Os módulos de dados e de cartografia do GIPAF estão integrados, sendo as coordenadas geográficas um dos pontos em comum. Sobre a cartografia digital, georeferenciada, foram distribuídos geograficamente os arqueossítios, utilizando uma simbologia própria, de acordo com o tipo de sítio (Figura 4).

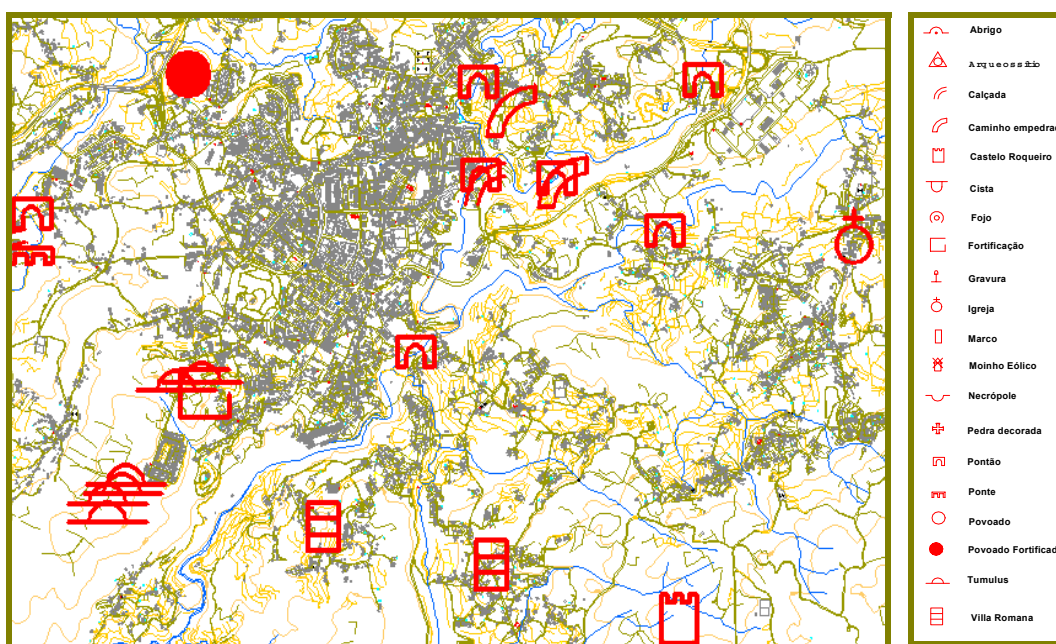


Figura 4 – Carta Arqueológica do Concelho de Fafe (vista parcial)

Cada sítio arqueológico foi ainda associado a um tipo de zona, definida como zona cautelar ou zona *non edificanti*, para as quais foram estabelecidos zonamentos conforme o terreno.

Esta base permitirá às entidades competentes, criar/complementar um Sistema de Informação Geográfica, onde para além da distribuição espacial dos sítios, se pode estabelecer a ligação à Base de Dados para visualização de outras informações/imagens a eles associada.

## 2.3. BANCO DE IMAGENS

As alterações constantes verificadas no território, nomeadamente as resultantes da forte pressão urbanística aumentam a importância da existência de Sistemas de Informação, onde a componente de imagem esteja presente e associada aos dados.

Se a possibilidade de visualizar imagens dos sítios parece, desde logo, uma componente fundamental do Sistema de Informação, esse valor crescerá se considerarmos o facto dos sítios poderem ser destruídos total ou parcialmente, de forma súbita ou inesperada. Alguns dos arqueossítios referidos no inventário arqueológico de Fafe de 1983 não foram localizados. A associação de imagens do sítio e de panorâmicas do local poderá constituir um instrumento fundamental na sua localização.

Da inventariação de sítios de património arqueológico no concelho de Fafe consta um extenso banco de imagens, que foram devidamente catalogadas e classificadas.

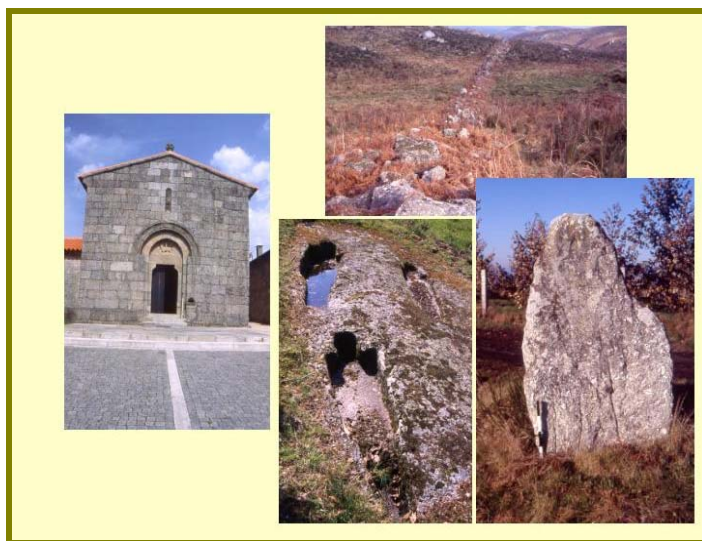


Figura 5 – Fotos de arqueossítios do Concelho de Fafe

### **3. SISTEMA DE APOIO À PROSPECÇÃO ARQUEOLÓGICA**

Foi já referida a importância das Tecnologias de Informação por permitirem a gestão dos dados do património, para que os investigadores possam trabalhá-los e as entidades responsáveis assegurem a sua conservação e valorização. A elaboração das cartas patrimoniais, com base no inventário de todos os bens prospectados, fomenta também a



cartografia de zonas de risco que devem ser preservadas, por forma a não danificar património existente.

No entanto, muitas vezes as intervenções no subsolo vêm revelar a existência de património desconhecido e ainda não inventariado. Dada a grande pressão construtiva, nomeadamente em meios urbanos, o património é descoberto não como resultado de sondagens e escavações programadas mas, em muitos casos, em consequência de empreendimentos públicos e privados cujos impactos patrimoniais são, muitas vezes, subavaliados.

Existindo já um inventário dos sítios arqueológicos é possível utilizar a Descoberta de Conhecimento em Bases de Dados (DCBD), para apoiar o desenvolvimento de algum conhecimento sobre o património oculto, tendo em vista a sua preservação e mesmo posterior valorização. Por outro lado a avaliação dos impactos poderia ser mais precisa se, antes de qualquer intervenção no subsolo, fosse possível utilizar uma ferramenta de apoio indicativa do potencial património de determinada zona ou local.

O modelo preditivo apresentado a seguir, Sistema de Apoio à Prospecção Arqueológica (SAPA), pretende ser um contributo para este objectivo, ao desenvolver um sistema para apoiar o arqueólogo e as instituições da tutela na prospecção e cartografia de potenciais arqueossítios.

O SAPA tem por base principal um inventário realizado sobre o património arqueológico de Trás-os-Montes Oriental, reunida em catálogo [Lemos 1993] e utiliza a ferramenta de Descoberta de Conhecimento *Clementine v5.2 do SPSS Inc.*

No processo de DCBD realizam-se várias tarefas, dedicando-se as primeiras à preparação dos dados, sobre os quais se vão aplicar técnicas de DM, após o que se segue a validação e incorporação de conhecimento já existente. Todo o processo é muito iterativo (figura 6), pelo que no final de cada tarefa pode existir a necessidade de voltar a realizar tarefas anteriores, para incluir alterações identificadas em fases mais avançadas.

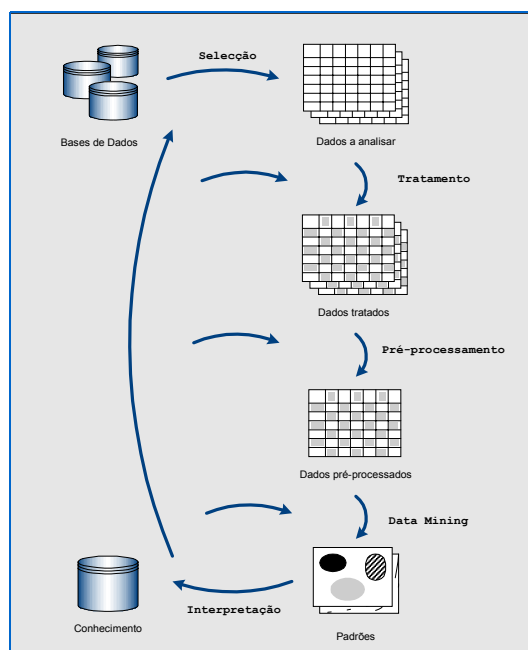


Figura 6– Fases do processo de DCBD (Adaptado de [FAYYAD *et al.* 1996])

Na figura 6 apresentam-se as 5 etapas que integram o processo de descoberta de conhecimento:

- Seleção dos dados;
- Tratamento dos dados;
- Pré-processamento dos dados;
- Aplicação de algoritmos de Data Mining;
- Interpretação e validação dos resultados.

O modelo preditivo SAPA foi desenvolvido executando todas estas tarefas, algumas delas repetidas várias vezes, para melhorar o seu desempenho.

### 3.1. SELECÇÃO DE DADOS

Na fase de **selecção dos dados** verificaram-se os dados disponíveis, representados nas tabelas da figura 6. Sobre as tabelas de dados iniciais foi feita uma selecção, que constituiu a primeira fase do processo de DCBD.

Tendo em mente o objectivo do trabalho, começou-se por retirar todos os dados que, por terem carácter meramente informativo, não são relevantes para este processo. Foi o caso das **Referências Bibliográficas**, **Topónimos** dos sítios, **Número** atribuído no catálogo, código de divisão administrativa (**LOCADM**), **Lugar** e número da **Carta Militar**.

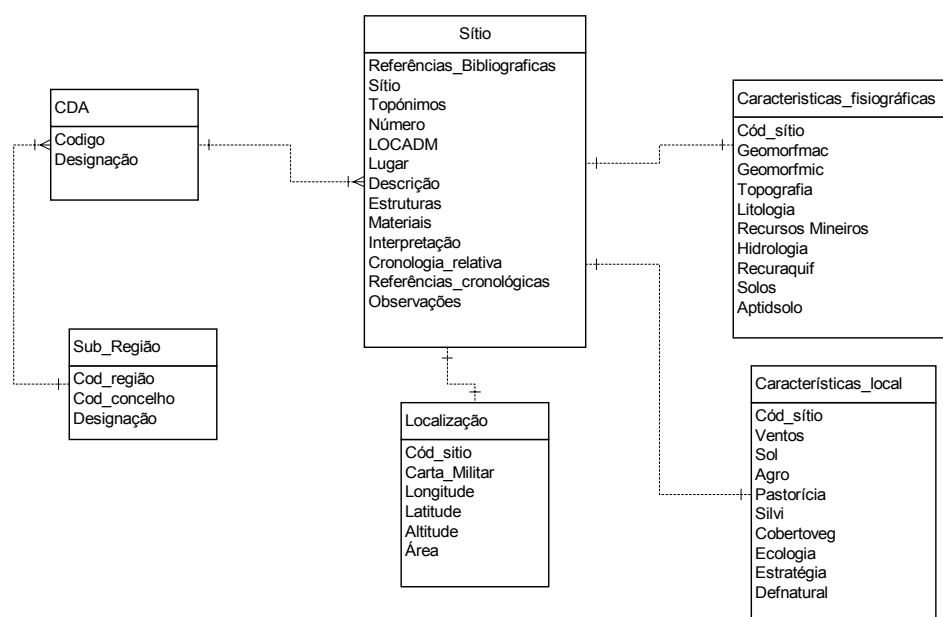


Figura 7 – Componentes da Base de Dados de sítios arqueológicos [Lemos 1993]

Aplicada a **selecção dos dados**, de acordo com o tipo de informação contida nos diversos atributos, procedeu-se à avaliação dos valores dos campos seleccionados.

### 3.2. TRATAMENTO DOS DADOS

O **tratamento dos dados** resultantes da fase de selecção iniciou-se com o processo de limpeza dos mesmos, tendo sido dada especial atenção ao tratamento dos valores omissos e dados inconsistentes ou inválidos. Os valores em falta para **altitude**, **geomorfologia**, **topografia** e **tipo de solos**, foram preenchidos sempre que identificáveis pelas as respectivas coordenadas geográficas.

Considerando que os atributos como a **Exposição solar**, **Ventos**, **Agricultura**, **Pastorícia**, **Silvicultura**, **Cobertura vegetal**, **Recursos mineiros**, **Estratégia** e **Defesa natural**, registavam uma taxa elevada de valores em falta e, não sendo possível

preencher os respectivos campos, optou-se por retirá-los do conjunto de dados a analisar. Utilizando ferramentas de manuseamento de dados, foram encontrados e também retirados todos os registos com informação em duplicado e que resultaram da junção de Bases de Dados distintas, onde alguns sítios arqueológicos foram objecto de múltiplos tratamentos.

Recorrendo às ferramentas gráficas disponibilizadas pelo *Clementine*, como os histogramas e gráficos de distribuição, foi possível visualizar os dados e identificar algumas inconsistências e erros de digitação.

### **3.3. PRÉ-PROCESSAMENTO DOS DADOS**

Concluídas as duas fases anteriores é feito um **pré-processamento dos dados**, de forma a facilitar a sua análise. As principais transformações operadas sobre os dados de sítios arqueológicos centraram-se na eliminação de variáveis correlacionadas e na análise e normalização de variáveis com significância no modelo final.

Valores contínuos da **Altitude**, **Longitude** e **Latitude** foram transformados em valores discretos, para serem analisados e agrupados por classes (Figura 8).

O campo **Tipologia**, ilustrado no gráfico da figura 8, resultou do pré-processamento realizado aos campos **Interpretação**, **Estruturas** e **Materiais** existentes na Base de Dados inicial. Estes campos continham um pequeno texto com a descrição do tipo de sítio identificado e dos materiais lá encontrados. A cada descritivo fez-se corresponder uma ou mais classes criadas para o efeito e foi registada a equivalência com um valor do campo **Tipologia**.

As classes criadas para o campo **Tipologia** apresentam os valores de “povoados”, “fortificações”, “povoados fortificados”, “necrópoles”, “santuários”, “rede viária”, “arte rupestre”, “epigrafia”, “minas”, “esconderijos” e “tesouros”.

Durante a fase de pré-processamento dos dados verificou-se ainda que alguns registos continham múltiplos valores para **Estruturas**, **Interpretação** e **Cronologias**. A título de exemplo refere-se o caso de registos com dados como “*vicus*; necrópole”, ou

“tesouro monetário; habitat romano” que, aparecendo na coluna **Interpretação**, correspondem cada um a duas classes de **Tipologia**. Fazendo a correspondência com as classes criadas para agrupamento e normalização dos dados, os valores de **Interpretação** “*Vicus*; necrópole” correspondem à **Tipologia** “povoado” e “necrópole” e os valores como “tesouro monetário; habitat romano” correspondem à **Tipologia** “tesouro” e “povoado”.

Os *habitats* foram muitas vezes locais de povoamento durante vários períodos cronológicos distintos, sendo frequente encontrar para um mesmo local duas cronologias distintas. Para estes casos procedeu-se também ao desdobramento dos registos.

Os recursos aquíferos, certamente relevantes na escolha dos locais de povoamento, estavam caracterizados nas tabelas de inventário nas colunas **Hidrologia** e **Recursos aquíferos**. Para facilitar o tratamento e análise da informação, criou-se uma nova coluna – **Hierarquia hidrográfica**, que generaliza os dados relativos, criando classes de 1 a 6 de acordo com a maior ou menor proximidade dos cursos de água principais (Figura 8).

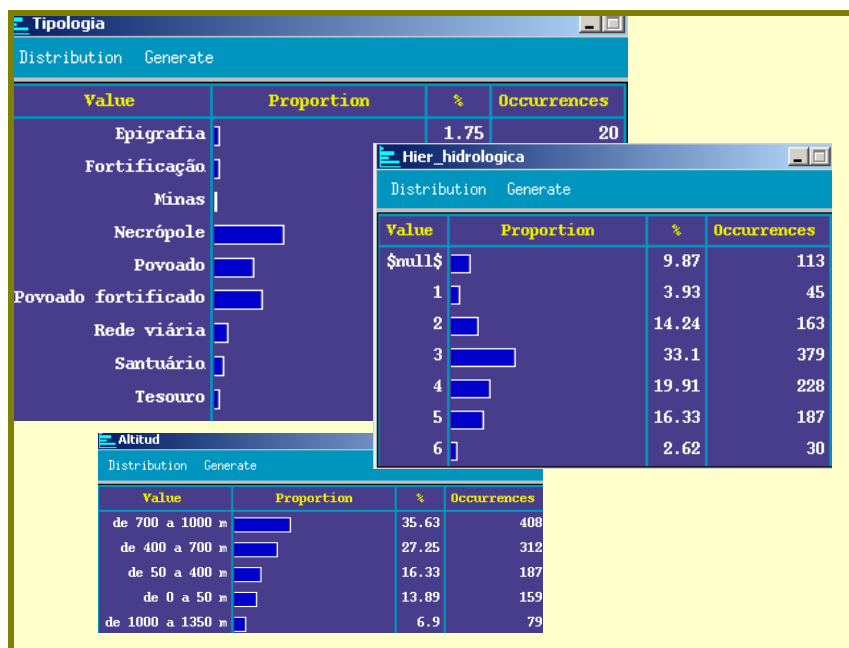


Figura 8– Classes criadas para os valores da **Altitude**, **Tipologia** e **Hierarquia hidrográfica**

Um outro gráfico apresentado na Figura 8 caracteriza a distribuição dos arqueossítios, de acordo com a tipologia criada. Assim pode verificar-se que a maioria dos arqueossítios são “necrópoles”, “povoados” e “povoados fortificados”. Poderá, no entanto, causar alguma surpresa o facto de existirem mais “necrópoles” do que “povoados”, uma vez que seria expectável que as necrópoles estivessem associadas a povoados. Tal acontece porque foram identificadas várias “necrópoles megalíticas”, tendo sido reduzidas as localizações dos respectivos “povoados”.

A Base de Dados de trabalho (Figura 2) contém alguns campos com informação correlacionada, isto é com alguma informação redundante no contexto deste trabalho. É o caso da coluna **Referências cronológicas** que, para a DCBD, introduz informação correlacionada com a coluna **Cronologia**, pelo que foi suprimida. As colunas **Hidrologia** e **Recursos aquíferos** apresentam também informação altamente correlacionada com a coluna **Hierarquia hidrográfica**, criada para generalização dos dados sobre a rede hidrográfica. Deste modo, as colunas que lhe deram origem, **Hidrologia** e **Recursos aquíferos** foram retiradas. Retiraram-se também as colunas **Interpretação**, **Estruturas** e **Materiais** por estarem altamente relacionadas com os dados da **Tipologia**, que resultaram da sua normalização e generalização.

Como resultado das operações realizadas, a tabela de dados a ser utilizada nas tarefas de Data Mining, ficou reduzida aos campos representados na Tabela 1.

<b>Sítios arqueológicos</b>
Tipologia
Cronologia
Altitude
Latitude
Longitude
Litologia
Geomorfologia
Geomorfologia_mic
Topografia
Hierarquia hidrográfica
Solos
Paisagem

Tabela 1 – Tabela de dados após o pré-processamento

### 3.4. ANÁLISE DE RELAÇÕES ENTRE OS DADOS

A partir dos elementos da Tabela 1 iniciou-se a fase de exploração dos dados, utilizando algumas técnicas de visualização disponíveis no *Clementine*, nomeadamente os *Web Nodes*. Esta forma de apresentação gráfica proporciona uma interpretação dos relacionamentos existentes entre os dados e possibilita a identificação das variáveis mais influentes na construção do modelo preditivo.

Este tipo de gráficos facilitam a identificação de algumas relações interessantes entre dois ou mais atributos simbólicos. As ligações são expressas graficamente, através de pontos, linhas e linhas sombreadas. As relações mais fortes são as desenhadas a traço contínuo mais carregado, passando a tracejada quando estamos perante relações fracas. Dados não ligados indicam que não foi identificada qualquer relação entre eles.

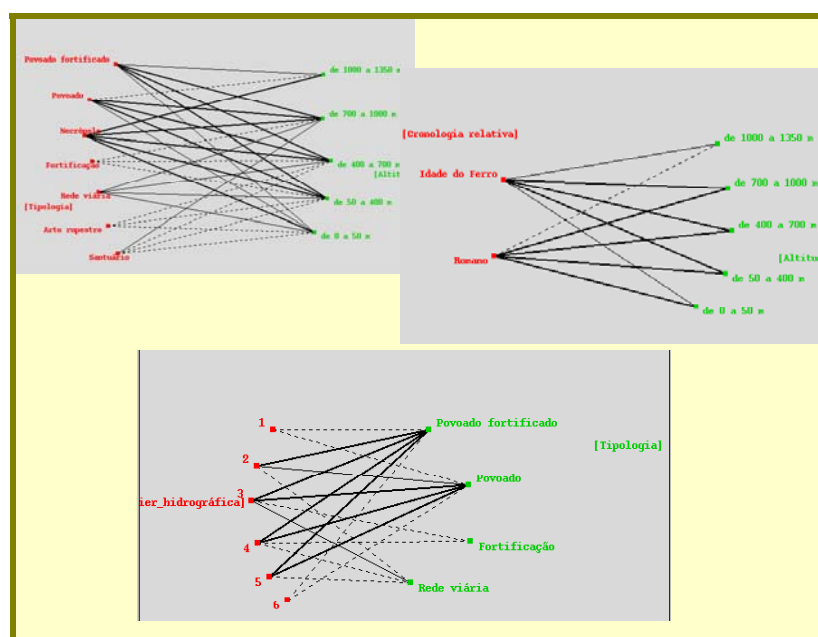


Figura 9 – *Web nodes* que relacionam a **Tipologia** com a **Altitude e Hierarquia hidrográfica e a Cronologia com a Altitude**

Fazendo uma breve análise aos gráficos da Figura 9 pode constatar-se que os povoados, povoados fortificados e as necrópoles se distribuem por todo o território da região. No entanto, o gráfico das cronologias permite detectar que no período romano foram abandonados alguns povoados da Idade do Ferro, situados em altitudes acima dos 1000

m. Tal facto deve-se a uma estratégia diferenciada de povoamento, nestes dois períodos cronológicos. Na Idade do Ferro há uma procura de locais com boas condições de visibilidade e boas características defensivas, havendo um equilíbrio com o contexto natural, enquanto que no período romano houve uma hierarquia de povoamento influenciada por uma nova economia e uma nova rede de caminhos.

Curiosamente as “fortificações” aparecem a altitudes intermédias e os “povoados fortificados” estão distribuídos por todo o território. De facto, as “fortificações” isoladas ou associadas a “povoamentos” não privilegiam apenas a “altitude”, podendo concluir-se que as “fortificações” tinham um carácter multifuncional, servindo estratégias de defesas, de controlo e delimitação do território envolvente ou até como símbolo arquitectónico.

A rede viária não se distribui por todos os patamares de “altitude”, o que confirma o conhecimento de que os engenheiros romanos evitavam não só as cotas muito elevadas, como também as grandes variações de altitude. O facto de haver “povoados” e “povoados fortificados” acima dos 1000 m, sem que a “rede viária” seja localizada nesse patamar de “altitude” deve-se ao traçado das vias não acompanhar a distribuição dos “povoados”, mas obedecer a um plano lógico e a um estudo prévio da geomorfologia.

O gráfico da Figura 9 estabelece também as relações entre as **Tipologias** e a **Hierarquia hidrográfica** e, tal como já se tinha verificado no gráfico da Figura 8, os locais mais próximos e mais afastados dos cursos de água principais (1 e 6 na hierarquia hidrográfica) são os menos ocupados pelos “povoados”, “fortificações” e “povoados fortificados”. O facto dos locais mais próximos do curso de água primário da região, que para a maioria dos sítios é o rio Douro, se caracterizarem por terem poucas **Tipologias** localizadas poderá ter várias explicações. Uma delas pode atribuir-se à circunstância deste rio ter margens com vertentes escarpadas e de difícil acessibilidade, factor que poderia ter impedido uma ocupação mais densa. Outra justificação está relacionada com a prospecção pouco intensiva destas áreas, no âmbito de projectos de Arqueologia.



### 3.5. APLICAÇÃO DE ALGORITMOS DE DATA MINING

Terminadas as fases de compreensão dos dados, selecção, tratamento e pré-processamento dos mesmos, procedeu-se à aplicação de técnicas de Data Mining, para construção do modelo preditivo de arqueossítios. Os dados foram divididos em dois conjuntos. Um desses conjuntos, normalmente o mais pequeno, é o conjunto de treino e o outro é o conjunto de testes. O primeiro irá ser usado para construir o modelo, servindo o outro para o testar e validar.

Dado o carácter previsionial do modelo que se pretende construir aplicou-se ao conjunto de treino um algoritmo de indução de árvores de decisão (Algoritmo C5.0) e, posteriormente, um algoritmo de redes neuronais [Berry e Linoff 2000].

Da aplicação do algoritmo C5.0 aos dados de treino resultaram algumas regras que se apresentam na Figura 10.

```
Cronologia relativa Romano
Geomorfmic [Alto Chá Cume] -> Epigrafia
Geomorfmic Alvêolo -> Povoado
Geomorfmic Arriba -> Vestigios
Geomorfmic Castelo -> Vestigios
Geomorfmic Inselberg -> Povoado
Geomorfmic Ladeira -> Rede viária
Geomorfmic Lombeiro -> Povoado
Geomorfmic Margem -> Povoado
Geomorfmic Monte -> Minas
Geomorfmic Outeiro -> Povoado
Geomorfmic Terraço -> Povoado
Geomorfmic Vertente -> Rede viária
Geomorfmic Esporão
  Altitud [de 0 a 50 m' de 1000 a 1350 m'] -> Santuário
  Altitud de 50 a 450 m -> Santuário
  Altitud de 700 a 1000 m -> Fortificação
  Altitud de 450 a 700 m
    Hier hidrográfica [null$ 1 3 4 6] -> Povoado
    Hier hidrográfica 2 -> Povoado
    Hier hidrográfica 5 -> Santuário
Geomorfmic Planalto
```

Figura 10 –Regras obtidas pela aplicação do algoritmo C5.0

As regras produzidas pelos algoritmos de árvores de decisão são escritas numa linguagem natural, o que facilita bastante a sua leitura e interpretação pelos especialistas da área.

Pode ler-se nas regras obtidas (Figura 10) que o primeiro critério utilizado para classificar o tipo de sítio arqueológico é a geomorfologia e, quando necessário, é usada a altitude e/ou hierarquia hidrográfica.

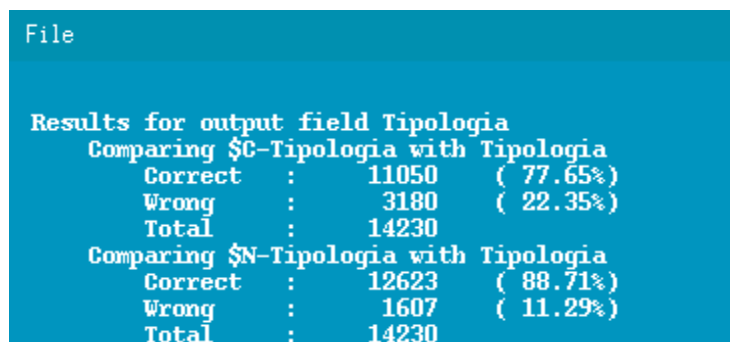
Por exemplo, para a geomorfologia esporão, isto é, um relevo secundário que sobressai de um conjunto mais vasto, é frequente encontrarem-se Povoados, Fortificações ou

Santuários. Neste caso a geomorfologia não é suficiente para caracterizar a tipologia, pelo que se recorre a outras características. Para a geomorfologia esporão, nos locais de altitudes mais baixas (abaixo de 450 m) ou mais elevadas (acima dos 1000 m) encontram-se Santuários, enquanto que nos patamares de altitude intermédios se podem encontrar Povoados e Fortificações.

Estas regras estão de acordo com as descobertas arqueológicas já realizadas, onde é frequente encontrarem-se Santuários no cimo dos montes, associados à divinização dos mesmos. Cita-se, a título de exemplo, o santuário do Deus Larouco, onde foi encontrada uma ara votiva evocando monte Larouco. Em altitudes mais baixas, na proximidade dos rios, é também frequente encontrarem-se Santuários, de que são exemplo os grandes santuários de arte rupestre encontrados junto dos cursos de água do Douro, Tejo ou Guadiana.

A validação do modelo obtido é feita através da confrontação dos resultados com o conhecimento arqueológico existente e também pela sua aplicação ao conjunto de testes.

Apresenta-se na Figura 11 a avaliação qualitativa do modelo preditivo em estudo, resultante da aplicação do algoritmo de indução de regras (\$C) e da rede neuronal (\$N) aos dados de teste.



```
File
Results for output field Tipologia
Comparing $C-Tipologia with Tipologia
Correct : 11050 ( 77.65%)
Wrong : 3180 ( 22.35%)
Total : 14230
Comparing $N-Tipologia with Tipologia
Correct : 12623 ( 88.71%)
Wrong : 1607 ( 11.29%)
Total : 14230
```

Figura 11 – Análise qualitativa da aplicação do modelo ao conjunto de testes.

## CONCLUSÃO

Ao longo dos anos as sociedades humanas adoptaram estratégias diferenciadas em relação ao espaço e ao meio ambiente. O homem passou de uma postura recolectora e de aproveitamento dos recursos naturais para atitudes mais intervencionistas que têm causado grande impacto na paisagem. No entanto, embora bastante desfasada da industrialização e da urbanização tem-se desenvolvido também uma atitude conservacionista dos valores patrimoniais. O património arqueológico, por ser um dos recursos culturais mais sensíveis, uma vez que é um recurso não renovável, tem sido objecto de algumas acções de salvaguarda, nomeadamente através da adopção de medidas legislativas, criação de museus e de parques culturais. Desenvolveu-se uma nova indústria baseada na criação de roteiros culturais, recreação de espaços e actividades de épocas passadas e na divulgação de monumentos.

No sistema legislativo e nas recomendações, nomeadamente da Convenção Europeia, ressaltam algumas preocupações quanto à necessidade de se proceder à inventariação dos valores arqueológicos e da inserção dos sítios registados no planeamento urbano e do território. No entanto, ainda recentemente o estudo sobre as barreiras à produtividade encontradas em Portugal, realizado pelo McKinsey Global Institute e encomendado pelo Governo Português, encontramos o ordenamento do território como uma dos factores burocráticos que contribuem para a nossa baixa produtividade.

Muito há a fazer para preservar e valorizar o nosso património e realizar um bom ordenamento do território, de acordo com vectores estruturais que contemplem todos os factores relevantes a esta actividade. O contributo das Tecnologias de Informação tem-se mostrado relevante, permitindo gerir a informação existente e colocá-la ao serviço da investigação e da divulgação cultural. Abre ainda algumas oportunidades na criação de conhecimento arqueológico que vai além da inventariação dos bens e da criação de itinerários turísticos ou pedagógicos, bem como a delimitação de zonas de protecção do património inventariado.

O conhecimento gerado com base no saber já existente e nos inventários construídos

poderá levar-nos um pouco mais além no conhecimento das matrizes de povoamento existentes no passado, dos critérios a que obedeceriam e dos fenómenos de mudança e continuidade [Lemos 1987].

Só com um conhecimento profundo do passado poderemos estabelecer uma política de conservação, cujos critérios não entrem em ruptura com o sentido histórico da paisagem.

## REFERÊNCIAS BIBLIOGRÁFICAS

Berry, Michael e Gordon Linoff, “*Mastering Data Mining – The Art and Science of Customer Relationship Management*”, Wiley Computer Publishing, New York, USA, 2000.

Fayyad, U. M., G. Piatetsky-Shapiro, e R. Uthurusamy (Eds.), “*Advances in Knowledge Discovery and Data Mining*”. The MIT Press, Massachusetts, 1996.

Han, Jiawei e Micheline Kamber, “*Data Mining: Concepts and Techniques*”, Morgan Kaufmann Publishers, 2001.

Lemos, Francisco Sande, “*Arqueologia e Território*”, Universidade do Minho, Actas do I Congresso Luso-Galego de Conservação e Ambiente, Braga, 1987.

Lemos, Francisco Sande, “*Povoamento Romano de Trás-os-Montes Oriental*”, Universidade do Minho, 1993, Tese de Doutoramento.

Lemos, Francisco Sande, “*História, Arqueologia e Organização do Território*”, Universidade do Minho, Actas do Seminário Pensar o Ordenamento do Território – Ideias, Planos, Estratégias, Lisboa, 2000.

Santos, Maribel Yasmina Campos Alves, “*PADRÃO – Um sistema de Descoberta de Conhecimento em Bases de Dados Georeferenciadas*”, Universidade do Minho, 2001, Tese de Doutoramento.

Rodrigues, Maria de Fátima, Carlos Ramos e Pedro Rangel Henriques, “*Extracção de Conhecimento em Sistemas de Informação Imprecisos*”, EEI’98, 1998.

SPSS, Clementine, User Guide, Versão 5.2, SPSS Inc., 1999.