

# Base de dados terminológica

**Proponente** Gabriel David  
FEUP  
gtd@fe.up.pt

---

---

Esta proposta especifica um módulo assente numa base de dados terminológica vocacionada para sectores específicos com suporte para várias línguas, características dos termos configuráveis e possibilidade de definir associações entre termos. As interfaces de navegação e, desejavelmente, de manipulação deverão estar baseadas em browsers Web. A interoperabilidade com outros módulos e ferramentas externas será assegurada por uma interface XML, segundo as normas MARTIF (ISO 12200) e ISO12620.

---

---

**Contexto** Uma das actividades fundamentais no desenvolvimento de qualquer área do conhecimento é o estabelecimento de uma terminologia que defina os seus termos fundamentais. A terminologia só é útil se for amplamente conhecida e fácil de enriquecer. Propõe-se para isso o desenvolvimento de um módulo de software com acesso pela Web que permita armazenar a informação típica de uma terminologia.

As utilizações possíveis para uma base de dados terminológica são múltiplas, indo desde a consulta por interessados na área, à verificação ortográfica especializada ou ao suporte à tradução a variados níveis. Precisamente para permitir esta última utilização, pretende-se que seja possível representar o mesmo conceito usando termos de várias línguas. Como as ferramentas em que um módulo desta natureza pode ser útil são muitas é importante garantir a interoperabilidade. Para isso atender-se-á às normas existentes e utilizar-se-á a XML como veículo de troca de informação terminológica.

O desenvolvimento do projecto será apoiado pelo grupo de Tradução da Faculdade de Letras da Universidade do Porto.

# Especificações<sup>1</sup>

O sistema a desenvolver deve suportar múltiplas terminologias, à partida independentes (ex. Imunoterapia ou Informática Médica). Uma terminologia é composta por conceitos. Estes exprimem-se por termos (com uma ou mais palavras), havendo tipicamente um termo para cada língua suportada. Nalguns casos, serão admissíveis vários termos alternativos para a mesma língua (ex. “operating system”, “sistema operativo”, “sistema operacional”).

## Modelo de dados

Cada termo poderá ser descrito por várias características. Cada terminologia definirá as características utilizáveis, a respectiva cardinalidade (valor único ou múltiplo) e o seu carácter de obrigatoriedade. Será possível criar características novas mas de tipos predefinidos. Os tipos a considerar serão:

- Chave – que pode ser usado como referência, tipicamente um termo numa determinada língua (ex. “operating system”), indexado na BD.
- Atributo – característica cujos valores se encontram numa lista de valores predefinidos (ex. “Gramática: substantivo | adjetivo | verbo”). A língua poderia ser um atributo mas, pela sua importância no sistema, fica associada à chave.
- Texto – uma descrição livre (ex. “Definição”, “Contexto”, “Exemplo de uso”).
- Imagem – nome de um ficheiro que registe uma imagem alusiva ao termo em causa.
- Associação – referência a uma chave (ou conceito) existente na terminologia, segundo uma determinada perspectiva (ex. “Sinónimo”, “Relacionado”, “Não sinónimo”); as características deste tipo constituirão a base da navegação na terminologia.

A utilização, numa característica de tipo texto, de um termo que seja ele próprio uma chave deverá ser reconhecida pelo sistema e produzir uma associação automática da categoria “Referência”.

---

<sup>1</sup> Indicam-se entre parenteses rectos as funcionalidades de segunda prioridade.

Cada termo terá ainda associada meta-informação administrativa registando o criador dessa entrada na terminologia, a data de criação e da última alteração.

**Funcionalidade** O sistema apresentará um formulário de pesquisa de termos, devolvendo a página com o conceito respectivo, isto é, o conjunto de termos, nas várias línguas, que exprimem o mesmo conceito e as respectivas características. [As línguas e as características a apresentar devem poder ser configuradas para a sessão.] As referências existentes são materializadas por ligações para páginas congéneres.

As etiquetas que designam as características terão associada (*tooltip* ou ligação) uma explicação da sua razão de ser (ex. “Exemplo ≠ utilização do termo por autor referenciado”). [As etiquetas e outros textos da interface poderão estar numa língua escolhida pelo utilizador para a sessão (Português, Inglês)].

Para além da página de consulta, existirá uma página de edição de termos que permitirá criar novos conceitos, adicionar e alterar termos. As características admissíveis são as estabelecidas para a terminologia. O acesso a esta página deve ser sujeito a autorização e o autor e data das alterações à terminologia registado.

Existirá ainda uma página de edição de terminologias para criar, eliminar, exportar e importar terminologias. A exportação e a importação será feita em formato definido pelas normas MARTIF (ISO 12200) e ISO12620, tendo em atenção a especificação TBX (TermBase eXchange (TBX) format )do projecto SALT (Standards-based Access service to multilingual Lexicons and Terminologies).

[Finalmente haverá um mecanismo de autenticação/autorização de utilizadores, com os níveis (cada nível inclui os inferiores):

- o Gestão de terminologias (criação, eliminação, importação);
- o Gestão de termos (criação, eliminação e alteração de todos os conceitos e termos);
- o Edição de termos (criação, eliminação e alteração dos conceitos e termos de sua autoria);
- o Consulta e exportação.]

## Execução

O número de elementos do grupo de desenvolvimento do módulo deverá ser entre 2 e 4.

A tecnologia a utilizar para a base de dados, a interface, as operações e a validação de XML será objecto de decisão na primeira fase do desenvolvimento.