

## *Ontodomo*: um modelo de base de conhecimentos terminológicos

Joaquim Rodrigues Bento<sup>1</sup>

**RESUMO:** Neste artigo, apresentamos alguns resultados de nossa tese de Doutoramento, na qual propusemo-nos construir um modelo de *base de conhecimentos terminológicos* (BCT) a partir de um *corpus* especializado (*Domo corpus*). O objectivo principal deste estudo centrou-se na definição dos fundamentos teórico-metodológicos, assim como na experimentação de diferentes técnicas e sistemas de tratamento automático (e assistido) da língua, para a aquisição, a partir do *corpus*, da terminologia e das relações conceptuais necessárias à estruturação terminológica e modelização do *corpus* e do domínio. **Nestas breves páginas**, pretendemos apenas mostrar as características principais de nosso modelo BCT, implementado informaticamente no sistema *OntoDomo*.

**Palavras-chave:** Lexicologia; terminologia; terminologia descritiva; estruturação terminológica; Processamento de Linguagem Natural (PLN); Base de Conhecimentos Terminológicos (BCT); ontologia; recuperação de informação (RI).

**ABSTRACT:** In this article, we present some results of our Ph.D. thesis research in which we proposed to build a model of a *Terminological Knowledge Base (TKB)* based on a specialized *corpus* (*Domo corpus*). The main purpose of our research was the definition of the theoretical and methodological foundations, as well the experimentation of different Natural Language Processing (NLP) tools and techniques for the acquisition of terms and conceptual relationships, which are necessary both for the terminological structuring, the *corpus* and the domain modelling. **In these few pages**, we intend only to show the main features of our *TKB* model, implemented in an integrated computer system – *OntoDomo*.

**Keywords:** Lexicology; terminology; descriptive terminology; terminology structuration; Natural Language Processing (NLP); Terminological Knowledge Base (TKB); ontology; information retrieval (RI).

---

<sup>1</sup> Instituto Politécnico de Viseu, Portugal. E-mail: bentoesev@esev.ipv.pt.

## 1. Nota introdutória

**E**ste projecto, realizado no âmbito da investigação de Doutoramento (Bento, 2007) incidiu, como o nome pretende sugerir, no domínio de especialização tecnológico de *edifícios inteligentes e domótica*.

Nesta investigação, propusemo-nos construir um modelo de *base de conhecimentos terminológicos* (BCT), a partir de um *corpus* especializado, constituído por um conjunto de teses e artigos académicos, produzidos no âmbito da investigação de nível superior, em Portugal.

O objectivo deste estudo centrou-se na definição dos fundamentos teórico-metodológicos e na experimentação de diferentes técnicas e sistemas de tratamento automático (e assistido) da língua, para a aquisição, a partir do *corpus*, da *terminologia* e das *relações conceptuais* necessárias à *estruturação terminológica* e *modelização* do *corpus* e do domínio.

Remetendo toda a fundamentação para a obra supracitada, neste artigo, apresentamos a implementação informática, o sistema *OntoDomo*, desenvolvido no Centro de Linguística da Universidade Nova de Lisboa, “linha 2 – *Lexicologia, Lexicografia e Terminologia*”, sob a Direcção da Professora Doutora Teresa Lino, com os seguintes objectivos aplicacionais:

- A recuperação de informação e consulta assistida de documentação especializada.
- A navegação na estrutura conceptual do domínio (núcleo ontológico da BCT) e a visualização dos grafos conceptuais gerados pelo sistema;
- A consulta da informação terminográfica considerada relevante.

2. O modelo de dados *ontodomo*

A definição de um modelo de dados serve de orientação na análise do *corpus* e na aquisição dos dados conceptuais e terminológicos (cf. Condamines, 2003). Assim, a primeira tarefa consiste em definir o que é relevante *adquirir* no *corpus*. Neste estudo, adoptou-se o modelo de dados explicitado na figura 1, explanado através do conceito de *detector de gás*.

CONCEITO		82.00	DETECTOR DE GÁS
Domínio		003	protecção contra incêndio
Relação genérica	is_a	001	detector
Relações partitivas	part_of_c	002	sistema doméstico
	part_of_ph	003	
Relações funcionais	used_for	004	detecção de fugas de gás
	has_purpose	005	protecção
Relações específicas (exemplos)	entrada	006	presença de gás combustível
	saída	007	sinal eléctrico binário (0/1)
	método de detecção	008	

Figura 1: O modelo de dados *OntoDomo*: a entrada *conceito*

### 2.1. A Entrada Conceito

Para além do subdomínio, a entrada *conceito* inclui os seguintes elementos (Figura 1):

a) *A designação (normalizada) do conceito (label)*

A designação do conceito emerge da descrição do *corpus*, segundo uma metodologia ascendente. Da aquisição de *unidades terminológicas* e *relações conceptuais*, resulta um conjunto de unidades linguísticas designando o mesmo conceito, as quais, depois de estruturadas, foram validadas pelos especialistas, formando um paradigma linguístico, no exemplo, com os seguintes elementos:

detector de gás  
**detectores de gás**  
 detector de fuga de gás  
**detectores de fuga de gás**  
 detector de fugas de gás  
**detectores de fugas de gás**  
 sensor de gás  
**sensores de gás**  
 sensor de fuga de gás  
**sensores de fuga de gás**

Em seguida, foi estabelecida, pelos especialistas, a *cota de aceitabilidade terminológica* (*term acceptability rating*) segundo a classificação definida pelas normas ISO (cf. ISO 107-1:2000: 8):

- TP – termo privilegiado (*preferred term*);
- TA – termo admitido – tolerado (*admitted term*);
- TD – termo desaconselhado (*deprecated term*);
- TO – termo obsoleto (*obsolete term*).

No exemplo, *detector de gás* é assumido como o *termo privilegiado* (TP) e, conseqüentemente, é adoptado como *designação* (*label*) do respectivo conceito, partilhando a atribuição de um mesmo identificador numérico, neste caso, 82.00.

*b) A explicitação das relações genéricas*

A relação conceptual *genérica*, *is\_a*, relaciona cada conceito *subordinado* com o respectivo *conceito genérico*. A relação conceptual genérica constitui o travejamento de qualquer recurso ontológico baseado nos mecanismos da herança (cf. Bourigault e Aussenac-Gilles, 2004: 31).

*is\_a* (“detector de gás”, “detector”).

*c) A explicitação das relações partitivas*

Neste modelo, de acordo com a nossa pesquisa, são utilizadas duas das categorias da relação partitiva da taxinomia de *Winston et al.* (1987), consideradas mais relevantes, no domínio: a relação *componente-objecto* (*part\_of\_c*), e a relação *fase-processo* (*part\_of\_ph*).

*d) A explicitação das relações funcionais*

Entre as relações funcionais, «télicas», que têm sido objecto de aprofundada reflexão por parte de muitos autores como *Pustejovsky et al.* (2006) e, *e. g.*, *Vossen* (2006) e *Lautenbacher* (2000), utilizámos dois níveis de funcionalidade: a função imediata (*used\_for*) e as finalidades genéricas do um sistema domótico (*has\_purpose*).

*e) Explicitação das relações específicas*

Relativamente às relações específicas do domínio, sendo em grande número, limitámo-nos, neste estudo, ao tratamento, exemplificativo, de

algumas, que permitem a definição *diferencial* de sistemas conceptuais modelizados, como os de *sensor* e *detector*. Deve anotar-se que as aplicações de pesquisa, baseiam-se, essencialmente, na relação genérica (*is\_a*), afastando-se da tendência formal das ontologias e bases de conhecimentos, produzidas no âmbito da inteligência artificial.

## 2.2 A Entrada Termo

Como se explicita na figura 2, a componente linguística da base de conhecimentos terminológicos inclui a seguinte informação:

- A unidade terminológica (UT), designadora do conceito (termo privilegiado – TP), com o mesmo identificador numérico do conceito (82.00);
- Um exemplo (contexto) significativo, do uso da UT, no *corpus*;

<b>UT</b>	<b>82.00</b>	<b>DETECTOR DE GÁS</b>
<b>Categoria gramatical</b>	<b>01</b>	<b>s. m.</b>
<b>Definição terminográfica gerada</b>		<b>Detector</b> Entrada – presença de gás doméstico (butano, propano ou gás natural) Saída – sinal eléctrico binário (0/1) Usado para – detecção de fugas de gás Objectivo – protecção contra incêndios
<b>Exemplo</b>		[...] em caso de fuga de gás os <i>detectores de gás</i> permitem ao sistema actuar e fechar as electroválvulas de entrada de gás. Manuel SEVERIANO et al ( 2005 ) Page: 58j (1e occ.)
<b>UT, sinónimos e variantes do corpus</b>	<b>82.00</b> <b>82.01</b> <b>82.02</b> <b>82.03</b> <b>82.04</b>	<b>detector de gás</b> detectores de gás <b>detector de fuga de gás</b> detectores de fuga de gás <b>detector de fugas de gás</b> detectores de fugas de gás <b>sensor de gás</b> sensores de gás <b>sensor de fuga de gás</b> sensores de fuga de gás
<b>Variantes abreviadas</b>		
<b>Equivalente francês</b>		<b>détecteur de gaz</b>
<b>Equivalente inglês</b>		<b>gas detector</b>
<b>NOTAS</b>		

Figura 2: O modelo de dados *OntoDomo*: a entrada *termo*

– O conjunto dos *termos sinónimos* e suas *variantes*, no plano da expressão. As unidades sinónimas tomam os identificadores numéricos, incrementados em uma unidade, a partir do identificador da UT privilegiada: 82.01, 82.02, 82.03 e 82.04, tendo em vista a facilitação do tratamento informático. No exemplo de *detector de gás*, todos os termos sinónimos foram validados e classificados segundo a norma ISO (ISO 107-1:2000: 8). São também incluídas as variantes flexionais de número (plural).

- Variantes abreviadas (siglas e acrónimos);
- Termos equivalentes nas línguas inglesa e francesa;
- Um campo “NOTAS”.

Não existindo, como é reconhecido, praticamente, sinónimos absolutos, no plano do significado linguístico, sempre que a comutação dos termos *sinónimos* não altere o valor de verdade das proposições registadas no *corpus*, os termos são agrupados *numa única entrada terminológica*, sendo as restantes consideradas *variantes*, do plano da expressão, evitando-se, deste modo, a redundância resultante da multiplicação de “fichas terminológicas” desnecessárias. Realizado todo o trabalho de descrição do *corpus*, na fase de normalização linguística, a análise das propriedades dos objectos, quando possível, será também um critério a ter em conta (cf. Depecker, 2005: 11).

Uma lista alfabética de todos os termos, permitirá o acesso à *entrada terminológica* em que figura cada um dos termos, bem como à respectiva *entrada conceito*.

### 3. A multifuncionalidade da aplicação *ontodomo*

Não cabendo nestas páginas, por falta de espaço, uma descrição da arquitectura complexa do sistema *OntoDomo*, desenvolvido a par da presente investigação, no *Centro de Linguística da Universidade Nova de Lisboa*, importa referir que, em relação à interface com o utilizador (figura 3), a aplicação *OntoDomo* apresenta-se, no plano ergonómico, adequada ao utilizador não linguista / terminólogo, pensada, especialmente, para estudantes da área de domótica.



**Figura 3:** Ecrã de entrada da aplicação *OntoDomo*

Os utilizadores do sistema *OntoDomo* beneficiam do trabalho minucioso de análise linguística de *corpora* e do seu tratamento por especialistas de reconhecido prestígio, do domínio de *edifícios inteligentes e domótica* e, da área da *engenharia do conhecimento*.

### 3.1. Ontologia e Base de Conhecimentos Terminológicos: Edição e Navegação

O núcleo conceptual de uma base de conhecimentos terminológicos é constituído por uma ontologia. A navegação, numa ontologia permite, no plano didáctico e da ergonomia cognitiva, visualizar a estrutura de um sistema conceptual, reduzindo a multiplicidade das expressões linguísticas a um “mapa cognitivo” facilitador da compreensão do domínio, por exemplo, do microsistema conceptual de *detector*. Como evidencia a figura 4, o sistema.



Figura 4: Navegação no núcleo ontológico do sistema conceptual de *detector*; visualização dos conceitos específicos de *detector de movimento*

*OntoDomo* permite que o utilizador se desloque ao longo da taxinomia e visualize, através da interface gráfica, as relações genéricas existentes entre os conceitos da ontologia, podendo, a partir de qualquer conceito desta “espinha dorsal”, aceder aos pormenores da informação linguística contida na base de conhecimentos.

Neste modelo relacional, cada conceito define-se pelas relações que sustém em relação aos restantes, sendo a relação de *especialização*, (*is-a*) o *suporte* estrutural das restantes relações.

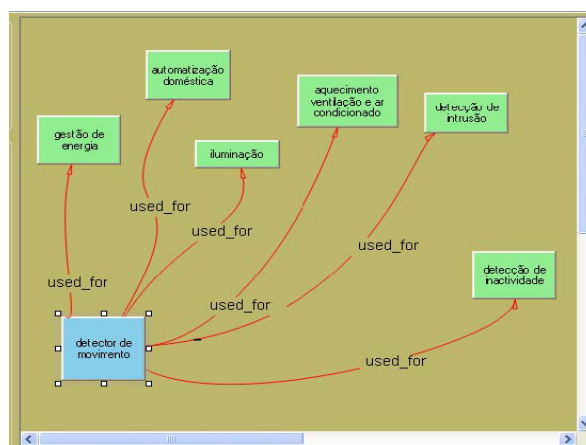


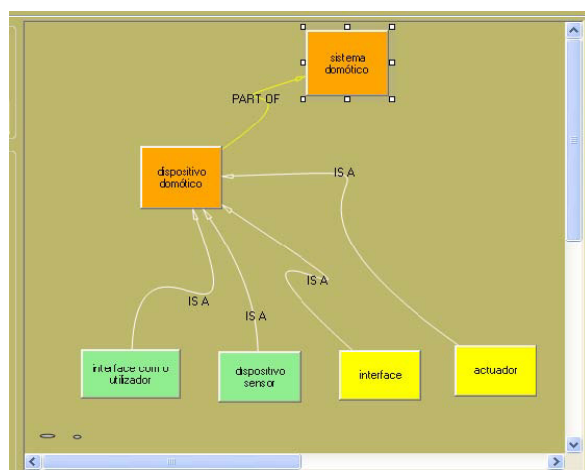
Figura 5: Grafo conceptual gerado pelo sistema *OntoDomo*, traduzindo o conjunto das asserções baseadas no predicado *used\_for*



Na figura 5, é representado um grafo conceptual, gerado pelo sistema *OntoDomo*, traduzindo as relações téticas do conceito *detector de movimento*, com os conceitos subordinados, específicos, de *função domótica: detecção de intrusão, gestão de energia, AVAC, iluminação, automatização doméstica e detecção de inactividade*.

Cada linha do grafo conceptual corresponde a uma *asserção lógica*, por exemplo, «*used\_for* (“*detector de movimento*”, “*detecção de intrusão*”)».

Na figura 6, são explicitadas relações partitivas e genéricas (*sistema misto de conceitos* (cf. (ISO 704, 2000: 14), sendo de anotar a herança de propriedades de *dispositivo domótico* pelos respectivos conceitos específicos: *dispositivo sensor*



**Figura 6.** Grafo conceptual gerado pelo sistema *OntoDomo*, representando um sistema misto de conceitos (relações *is\_a* e *part\_of\_c*)

(sensores e detectores), *actuador, interface e interface com o utilizador*. Na figura 7, está representada uma parte do *suporte* do conceito de *dispositivo sensor*, explicitados os conceitos subordinados de sensor: *sensor biométrico, sensor de luminosidade, sensor de temperatura e sensor de velocidade do vento* ou *anemómetro*.

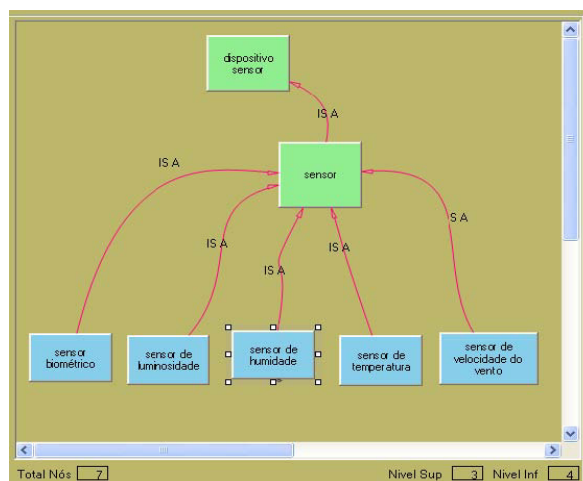


Figura 7: Suporte (parcial) da classe *dispositivo sensor*

### 3.2 Visualização da Informação Terminológica Contida na BCT

Como referido, a interface do *OntoDomo*, foi concebida para o utilizador não terminólogo, exibindo uma *ficha terminológica* em formato convencional, a partir da base de conhecimentos (figura 8). Para além da categoria gramatical, do domínio, das variantes em língua francesa e inglesa, é apresentado um contexto considerado significativo e uma definição, em língua natural, explicitando as características do conceito. Trata-se de um ambiente próximo do de um “dicionário terminológico”.



Figura 8: A exibição de conhecimentos terminológicos contidos na BCT

### 3.3. Um Motor de Pesquisa Assistida, Selectivo e Flexível

Um dos problemas dos actuais motores de pesquisa, como instrumentos da investigação, é o *ruído* resultante do uso de formas simples, e da existência de múltiplas variantes que o utilizador ou não retém na memória, por serem, frequentemente, muito numerosas, ou pelo tempo necessário à sua introdução manual, na entrada do motor de pesquisa. O método torna-se pouco eficaz, por excesso de *ruído*, ou de *silêncio*, representando grande desperdício de tempo, em particular, para o estudante inexperiente, que inicia um curso de especialização. Por outra parte, a selecção das palavras-chave não é devidamente instruída, resultando em *sobrecarga cognitiva*, defluente da obtenção de dados prolíferos, irrelevantes ou contraditórios. A presente investigação pretendeu dotar o sistema *OntoDomo* de conhecimentos suficientemente relevantes para que o mesmo possa constituir uma ferramenta verdadeiramente útil, na consulta de documentação especializada.

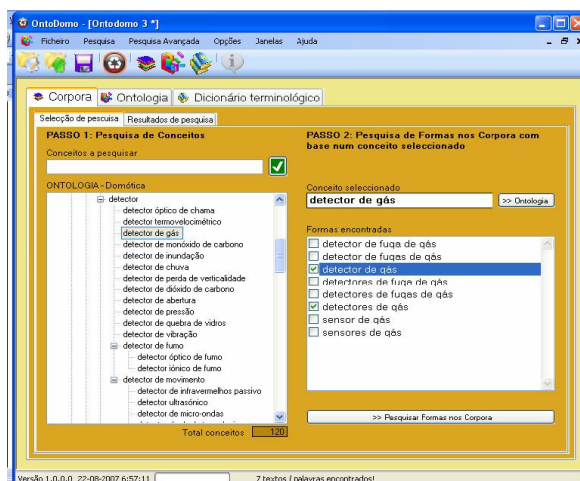


Figura 9: Pesquisa assistida, selectiva, a partir de um dos conceitos da ontologia

O princípio de funcionamento é ilustrado pela figura 9 e centra-se no núcleo conceptual da base de conhecimentos: *a ontologia*, a qual permite partir de uma síntese do domínio, cognitivamente acessível, para a diversidade das expressões linguísticas dos conceitos. Como ilustra a figura referida, o sistema *OntoDomo* propõe uma lista de termos ou variantes de um termo, bastando ao utilizador seleccionar todas ou apenas algumas das denominações sugeridas. Para o conceito *detector de gás*, o sistema propõe um conjunto de termos validados, extraídos do *corpus Domo*.

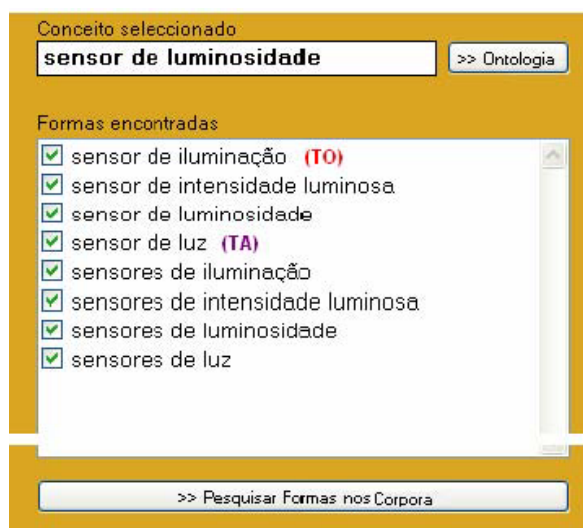
Propõe também a *designação privilegiada* do conceito (*detector de gás*), estabelecida pelos especialistas, no processo da normalização linguística. Por esta via, evita-se o *ruído* resultante da utilização de termos genéricos e beneficia-se de um trabalho aturado de análise linguística e modelização conceptual, tutelado pelos especialistas do domínio.

Como foi referido, oportunamente, a variação terminológica é um fenómeno normal, generalizado e que merece o maior destaque, em particular, na pesquisa e consulta de documentação técnica, assumindo dimensões que foram, durante muito tempo, subestimadas (cf., *e. g.*, Gaudin, 2003). A supervisão do especialista permitiu validar os dados introduzidos no sistema *OntoDomo*, tendo em conta que a observação minuciosa dos *corpora* suscita reflexão aprofundada, tanto no plano linguístico como no plano do domínio de especialização. Como exemplo, poderemos referir a substituição do termo *detector de gás* por *sensor de gás*, recolhendo o segundo cerca de 75% das ocorrências do conceito.

Observar o *corpus* e *reflectir* sobre os dados obtidos com os especialistas constitui um processo endógeno de valorização da língua portuguesa, como sistema modelizante da ciência e da tecnologia.

É por isso, oportuno que um sistema de pesquisa assistida possa oferecer alguma orientação aos estudantes que iniciam uma especialização. Esta também constitui uma valência funcional do modelo *OntoDomo*. Como se exemplifica através do tratamento do conceito *sensor de luminosidade* ou *sensor de intensidade luminosa*, o motor de pesquisa, para além de facultar as diversas designações, acrescenta informação visível sobre a classificação relativa às cotas de *aceitabilidade terminológica (CA)*, definidas pela *ISO*, por forma a que o utilizador seja prevenido quanto à valorização atribuída às diferentes opções.

A figura 10 ilustra esta função pedagógica do módulo de pesquisa do *OntoDomo*, ao classificar as diferentes unidades terminológicas. Deste modo o termo privilegiado, *sensor de luminosidade*, foi seleccionado para *label* do conceito: opção que implicitamente lhe confere o estatuto de termo preferido.



**Figura 10:** Proposta de pesquisa orientada através da indicação da *cota de aceitabilidade terminológica* (CA)

Dos restantes termos validados, um grupo apresenta-se normal, sem classificação, nem preferido nem desaconselhado. Os restantes são desvalorizados, sendo considerados termos *tolerados, desaconselhados ou obsoletos*. No exemplo, poderemos afirmar, com base no processo de validação, que tanto *sensor de luminosidade* como *sensor de intensidade luminosa* são termos valorizados, sendo quase arbitrária a selecção de um ou de outro como *label* do conceito. Em relação às designações *sensor de iluminação* e *sensor de luz*, a primeira foi classificada como *termo obsoleto* (TO); a segunda como *termo admitido / tolerado* (TA).

Esta informação de carácter normativo figura na base de conhecimentos terminológicos. Tratando-se de um *corpus* essencialmente académico, em geral, já passou pelo processo de avaliação académica. Por isso, são em número *muito reduzido* os termos não validados pelos especialistas. Em relação ao conceito *sensor de luminosidade*, foi eliminado, no processo de validação, o *candidato* a termo *sensor de iluminância*, sendo, por isso, excluído na base de conhecimentos não sendo, subsequentemente, proposto pelo sistema *OntoDomo*.

Um sistema de pesquisa deve ser eficaz, o que passa também pelo tratamento dos termos equivalentes do inglês, “importados” com a própria tecnologia. Um exemplo elucidativo provém da tecnologia domótica euro-

peia com maior implantação: a tecnologia *EIB* (*European Installation Bus*). Sendo uma tecnologia *não proprietária* (aberta), faculta e divulga as especificações técnicas, em várias línguas.

O termo *unidade de acoplamento ao barramento*, seria a tradução mais fiel de *bus coupling unit* (*BCU*). Contudo, no domínio da informática, *bus* constitui um empréstimo, sinónimo de *barramento* (cf. ILTEC: 1993). Assim, coexistem, no *corpus Domo*, como designações equivalentes a *bus coupling unit*, o termo acima referido e a denominação, mais curta, *acoplador de bus*. Contudo, o termo mais utilizado é a variante abreviada de *bus coupling unit*: *BCU*. Tendo em conta que a designação mais utilizada, no *corpus* é a sigla *BCU*, a mesma não pode deixar de figurar, como entrada essencial, do motor de pesquisa, como se ilustra na figura 11. Com efeito, o trabalho terminológico,

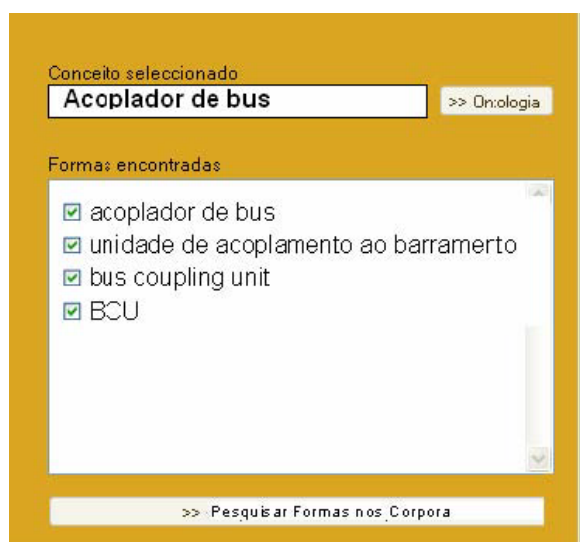


Figura 11: A inclusão de termos equivalentes

quando orientado para a recuperação de informação e consulta de documentação especializada, não pode prescindir dos meios linguísticos que conduzem às ocorrências dos conceitos, no *corpus*.

O termo equivalente não abreviado inglês, *bus coupling unit*, co-ocorre, quase sempre, com a sigla *BCU* e apresenta uma frequência reduzida, em contexto parentético, servindo para explicar a leitura da sigla:

```

03 5e| barramento ( designada por BCU - Bus Coupling Unit ) , e por um módulo
24 14a| are Elderly or Disabled BCU Bus Coupling Unit CSMA / CD Carrier S
24 54a| ste sistema é designado por BCU - Bus Coupling Unit ( unidade de acoplam
27 6a| FIGURA 3 - 17 - ESQUEMA DO BCU ( BUS COUPLING UNIT ) ... 47 FIGURA 3
27 7h| ASP - Active Server Page BCU - Bus Coupling Unit CAD - Computer Aide
27 73e| acoplamento ao barramento ( BCU - Bus Coupling Unit ) , acoplador de lin
28 21c| a unidade de acoplamento do bus ( Bus Coupling Unit - BCU ) , ilustrada
    
```

O mesmo não acontece com a sua abreviação, *BCU*, a qual apresenta um número de ocorrências tão elevado, no *corpus Domo*, que deve ser considerada um *empréstimo*, necessário à pesquisa, no *corpus*, como se infere da figura 12,

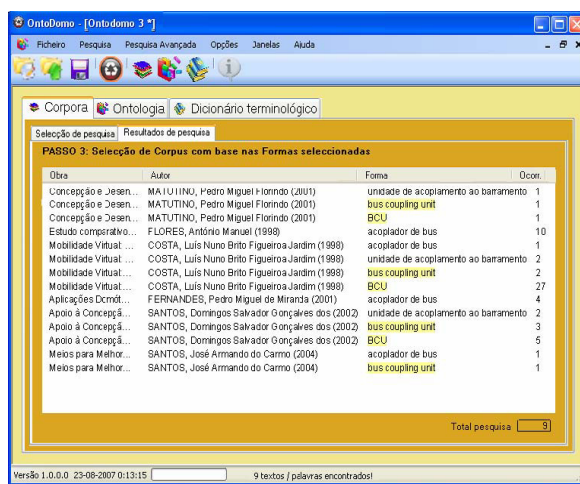


Figura 12: A primeira fase de acesso ao *corpus*, pelo sistema *OntoDomo*: o resumo dos dados

que sumaria os resultados da pesquisa solicitada na janela de entrada do *OntoDomo*. Observando os resultados facultados pelo *OntoDomo*, poderemos extrair as primeiras hipóteses para a prossecução da pesquisa, com base nas frequências:

```

acoplador de bus ..... 16
unidade de acoplamento ao barramento..... 5
bus coupling unit ..... 7
BCU ..... 33
    
```

No exemplo de *bus coupling unit*, verificamos que o termo não abreviado, co-ocorre sempre com a sigla *BCU*, não servindo, portanto, para recu-

perar mais informação. A informação que recupera é parcialmente redundante. O mesmo não acontece com a sigla *BCU*, indispensável ao motor de pesquisa, como revela a consulta em texto integral, facultada pelo *OntoDomo*. De resto, nos textos académicos, é frequente encontrar segmentos discursivos como o que sublinhámos na saída do *Ontodomo*:

A unidade de acoplamento ao barramento, *doravante designada por BCU*, é parte integrante de [...]

O que é peculiar, neste caso, é que a sigla *BCU* não corresponde ao termo não reduzido, *unidade de acoplamento ao barramento*, mas ao termo do inglês *bus coupling unit*.

Como pode observar-se, nos resultados da aplicação *Ontodomo*, todos os termos da pesquisa são, de imediato, assinalados a sombreado colorido. A forma particular da pesquisa, neste exemplo, *unidade de acoplamento ao barramento*, surge assinalada com uma cor mais escura.

A figura 13 representa o resultado de um experimento, ainda não completamente implementado no *OntoDomo*. Na versão actual, já estabilizada, apenas são propostas pela aplicação, para pesquisa, os termos em língua portuguesa e os empréstimos do inglês. Contudo, dada a convivência “indecisa”, frequente, de termos de língua portuguesa e inglesa, os equivalentes do inglês poderão vir a constituir também “variantes” a ter em conta, na instrução do motor de pesquisa.

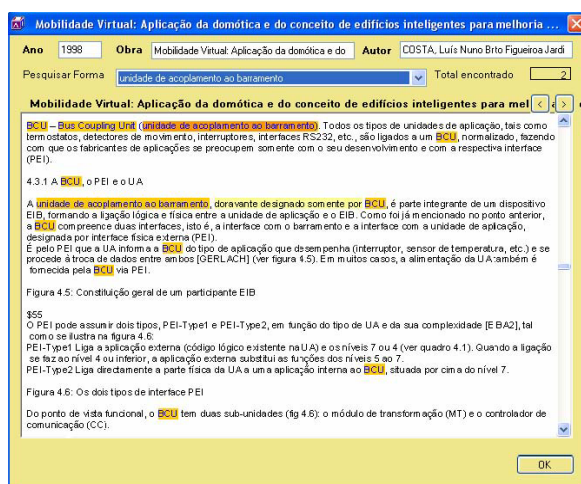


Figura 13: Os resultados da pesquisa assistida pelo *OntoDomo* do conceito *acoplador de bus*



Para além das principais funcionalidades já exemplificadas, o sistema *OntoDomo* permite também projectar no *corpus* todas as unidades terminológicas e suas variantes, assinalando (a cor azul) todos os termos do *corpus* que se encontram na base de conhecimentos, conforme ilustra a figura 14.

Este tratamento dos textos é, em si mesmo, uma forma de *assistir* a consulta (leitura) dos documentos. O nosso modelo de BCT permite, por esta via, a *indexação automática*, uma vez que a BCT forma, como o seu núcleo, a ontologia, um *índice estruturado*. Contudo, esta função ainda não foi implementada no sistema *OntoDomo*.

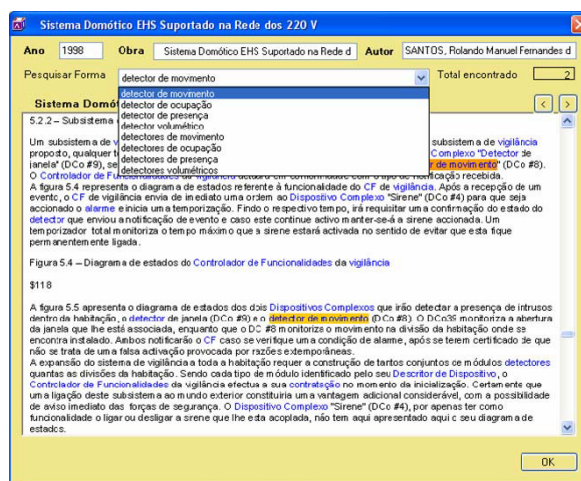


Figura 14: A projecção nos textos do *corpus* da terminologia do domínio

No estado actual de desenvolvimento, esta aplicação constitui, essencialmente, uma ferramenta para *consulta assistida de documentação especializada* no domínio da domótica. Para terminar estas referências à aplicação *OntoDomo*, cabe referir que a mesma constitui um sistema hipertextual avançado, potenciando a *navegação* entre a multiplicidade de elementos que articula, pertencentes aos *corpora* e à ontologia / BCT.

#### Referências bibliográficas

BACHIMONT, B. 2005. *Corpus et connaissances: De l'extraction linguistique à la modélisation conceptuelle*. In: CONDAMINES, A. (éd.). *Sémantique et corpus*. Paris: Hermès Science. p. 319-46.

- BANEYX, A. *et al.* 2005. Synergie entre analyse distributionnelle et patrons lexico-syntaxiques pour la construction d'ontologies différentielles. In : *Actes de la Conférence TIA 2005 (6èmes Rencontres Terminologie et Intelligence Artificielle)* Rouen: Université de Rouen, em linha, <http://www.loria.fr/~yannick/TIA2005/doc/>, consultado em 09/08/2005.
- BENTO, J. R. 2007. *Da constituição do corpus à construção de uma ontologia e base de conhecimentos terminológicos*. Dissertação (Doutoramento em Linguística / Lexicologia). Universidade Nova de Lisboa.
- BIÉBOW, B.; SZULMAN, S. 2000. Une approche terminologique pour catégoriser les concepts d'une ontologie. In: CHARLET, J.; ZACKLAD, M.; KASSEL, G.; BOURIGAULT, D. (éd.). *Ingénierie des connaissances: évolutions récentes et nouveaux défis*. Paris: Eyrolles, p. 325-36.
- BOURIGAULT, D. 2002. Upery: Un outil d'analyse distributionnelle étendue pour la construction d'ontologies à partir de *corpus*. In: *Actes de la 9ème Conférence Annuelle sur le Traitement Automatique des Langues (TALN 2002)*. p. 75-84.
- \_\_\_\_\_; CHARLET, J. 2000. Ontologies et textes. In : *Actes des 11èmes Journées Ingénierie des Connaissances (IC'2000)*. Toulouse: Centre pour l'Unesco, p. 7-8, em linha, <http://www.irit.fr/ic2000>, consultado em 03/04/2003.
- \_\_\_\_\_; AUSSENAC-GILLES, N.; CHARLET, J. 2004. Construction de ressources terminologiques ou ontologiques à partir de textes: un cadre unificateur pour trois études de cas. *Revue des Sciences et Technologies de l'Information (RSTI) série: Revue d'intelligence artificielle (RLA)* 18, n.º 1. 87-110.
- \_\_\_\_\_; JACQUEMIN, C.; L'HOMME, M. C. 2001. *Recent advances in computational terminology*. Amsterdam/ Philadelphia: John Benjamins.
- BUSA, F.; CALZOLARI, N.; LENCI, A. 2001. Generative lexicon and the SIMPLE model: Developing semantic resources for NLP. In: BUSA, F.; BOUILLON, P. (ed.). *The language of word meaning*. Cambridge: Cambridge University Press. p. 333-49.
- CABRÉ, M. T.; FELIU, J.; VIVALDI, J. 2004. Base de connaissances GENOMA: Le rôle de l'ontologie. *Workshop on Terminology, Ontology and Knowledge representation (TERMINO 2004)*. Lyon: Université Jean-Moulin Lyon 3. p. 19-25.
- CHEIN, M.; MUGNIER, M. 1992. Conceptual graphs: fundamental notions. *Revue d'Intelligence Artificielle* 6, n.º 4. 365-406.
- CHINGAREVA-SLAVINE, E. 2003. *Sémiotique linguistique et modélisation*. Lavoisier: Hermès Science.
- CONDAMINES, A. 2003. *Sémantique et corpus spécialisés: Constitution de bases de connaissances terminologiques*. Mémoire d'Habilitation à Diriger les Recherches. Toulouse: Universidade Toulouse 2.
- CONDAMINES, A. (éd.). 2005a. *Sémantique et corpus*. Lavoisier: Hermès Science.
- CORBIN, D. 1987. *Morphologie dérivationnelle et structuration du lexique*. 2 vols. Tubinga: Max Niemeyer Verlag.
- CORCHO, O. *et al.* 2003. Methodologies, tools and languages for building ontologies. Where is their meeting point?. *Data & Knowledge Engineering* 46. 41-64.
- CRUSE, D. A. 2000. *Meaning in Language. An introduction to Semantics and Pragmatics*. Nova York: Oxford University Press.
- DEPECKER, L. 2005. Contribution de la terminologie à la linguistique. *Langages* 157. *La terminologie: Nature et enjeux*. 6-13.

- FELIU, J.; VIVALDI, J.; CABRÉ, M. T. 2002b. Towards an ontology for a human genome knowledge base. *Third International Conference on Language Resources and Evaluation – LREC 2002*. Las Palmas: ELRA, p. 1885-1890, também disponível em linha, <https://nats-www.informatik.uni-hamburg.de/intern/proceedings/2002/LREC/htm/papers.htm>.
- FRADIN, B. 2003. *Nouvelles approches en morphologie*. Paris: PUF.
- GAUDIN, F. 2003. *Socioterminologie: une approche sociolinguistique de la terminologie*. Bruxelles: Duculot.
- GÓMEZ-PÉREZ, A. 1999. Développements récents en matière de conception, de maintenance et d'utilisation des ontologies. *Terminologies nouvelles* 19. 9-20.
- ISO 1087-1 2000-E/F. *Terminology – Vocabulary. Part 1: Theory and application. Travaux terminologiques. Vocabulaire Partie 1: Théorie et application*. Genève: Organisation Internationale de Normalisation.
- L'HOMME, M. 2002. Fonctions lexicales pour représenter les relations sémantiques entre termes. *Traitement automatique des langues* 43, n.º 1. 19-41
- LAUTENBACHER, O. 2000. *Les fondements perceptuels de la production de sens dans le lexique génératif*. Thèse (Doctorat). Université Marc Bloch de Strasbourg.
- LINO, M. T.; MEDINA, D.; MOREIRA, J. D.; CHICUNA, A. 2007. Rede de Neologia e de Terminologia em Língua Portuguesa (em situação de contacto de línguas). In: *XVII encontro da Associação das Universidades de Língua Portuguesa. AULP, Praia – Cabo Verde* (no prelo).
- \_\_\_\_\_. (éd.). 2003. *Mots et lexiculture, hommage à Robert Galisson*. Paris: Champion.
- MARRAFA, P. 2002. Portuguese WordNet: general architecture and internal semantic relations. *DELTA* 18. 131-46.
- MEYER, I. 2001. Extracting knowledge-rich contexts for terminography: a conceptual and methodological framework. In: BOURIGAULT, D.; JACQUEMIN, C.; L'HOMME, M. (ed.). *Recent advances in computational terminology*. Amsterdam/ Philadelphia: John Benjamins. p. 279-302.
- PUSTEJOVSKY, J. et al. 2006. Towards a generative lexical resource: the brandeis semantic ontology. In: CALZOLARI, N.; CHOUKRI, K.; GANGEMI, A.; MAEGAARD, B.; MARIANI, J.; ODJIK, J.; TAPIAS, D. (ed.). In: *Proceedings of the 5th International Conference on Language Resources and Evaluation. LREC 2006*, Génova: ELRA, p. 1702-05.
- SOWA, J. F. 2000. *Knowledge representation: logical, philosophical, and computational foundations*. New York: Brooks Cole Publishing.
- VOSSSEN, P. 2006. *EuroWordNet general document*, University of Amsterdam, em linha, <http://www.hum.uva.nl/~ewn>, consultado em 10/10/2008.
- WINSTON, M. E.; CHAFFIN, R.; HERRMANN, D. 1987. A taxonomy of part-whole relations. *Cognitive Science* 11, n.º 4. 417-44.
- XAVIER, M. F.; MATEUS, M. H. M. (ed.). 1990. *Dicionário de termos linguísticos*. v. I. Lisboa: Edições Cosmos.