

## Sistematização da terminologia da agricultura para o português brasileiro

*Tiago Pereira Rodrigues<sup>1</sup>*

*Gladis Maria de Barcellos Almeida<sup>1</sup>*

*Leandro Henrique Mendonça de Oliveira<sup>2</sup>*

*Ivo Pierozzi Jr.<sup>2</sup>*

O artigo segundo do Acordo Ortográfico (AO90) – assinado em Lisboa, em 16 de dezembro de 1990, por todos os países membros à época, e promulgado no Brasil em 29/9/2008 pelo Decreto nº 6.583 (no caso do Brasil, sua entrada em vigor está prevista para 1/1/2016, conforme o Decreto nº 7.875, de 27/12/2012) – determina que “os estados signatários tomarão, através das instituições e órgãos competentes, as providências necessárias com vista à elaboração, até 1 de janeiro de 1993, de um vocabulário ortográfico comum da língua portuguesa, tão completo quanto desejável e tão normalizador quanto possível, no que se refere às terminologias científicas e técnicas” (grifo nosso).

Nesse sentido, no âmbito do Projeto Vocabulário Ortográfico Comum da Língua Portuguesa (VOC), coordenado pelo Instituto Internacional da Língua Portuguesa (IILP), foi prevista a inclusão das primeiras terminologias, quais sejam: Agricultura, Pesca, Meio Ambiente, Saúde, Educação e Legislação.

Essa tarefa de sistematização terminológica foi iniciada pelo domínio da Agricultura por duas razões: primeiramente, já existe um vocabulário controlado bastante robusto, denominado Agrovoc, que vem sendo organizado pela Food and Agriculture Organization of the United Nations (FAO), desde a década de 1980. Esse vocabulário contém mais de 30.000 conceitos organizados de forma hierárquica e associativa com equivalentes em 26 línguas, dentre elas o português, variedade europeia. Em segundo lugar, o

---

<sup>1</sup> Universidade Federal de São Carlos - [tiago.rodrigues@colaborador.embrapa.br](mailto:tiago.rodrigues@colaborador.embrapa.br), [gladis@ufscar.com](mailto:gladis@ufscar.com)

<sup>2</sup> Embrapa Informática Agropecuária - [{leandro.oliveira, ivo.pierozzi}@embrapa.br](mailto:{leandro.oliveira, ivo.pierozzi}@embrapa.br)

GETerm/UFSCar tem um acordo de cooperação com a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), empresa vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (Mapa), criada em abril de 1973, que tem como missão “viabilizar soluções de pesquisa, desenvolvimento e inovação para a sustentabilidade da agricultura, em benefício da sociedade brasileira”<sup>3</sup>. Em razão de sua missão, a Embrapa detém ricas bases de conhecimento, vocabulários, corpora e, evidentemente, especialistas de domínios ligados à Agropecuária.

Este trabalho apresenta a tarefa de aproveitar o trabalho já feito pela FAO no Agrovoc, uma vez que esse vocabulário já traz os equivalentes em português europeu, e considerar as listas de termos obtidas de terminologias que estão sendo elaboradas atualmente pela Embrapa para alguns subdomínios da Agropecuária, tais como Intensificação Agropecuária, Agroecologia, Recursos Hídricos, Mudanças Climáticas, Informação Geoespacial, entre outros, para incluir a variedade brasileira do português no Agrovoc.

Para a tarefa de sistematização da terminologia brasileira da Agricultura e sua harmonização com a variedade europeia, foram desenvolvidas pela Embrapa: Informática Agropecuária uma interface denominada Agrovoc Search<sup>4</sup> e uma ferramenta denominada Etecam, que executa um matching de termos.

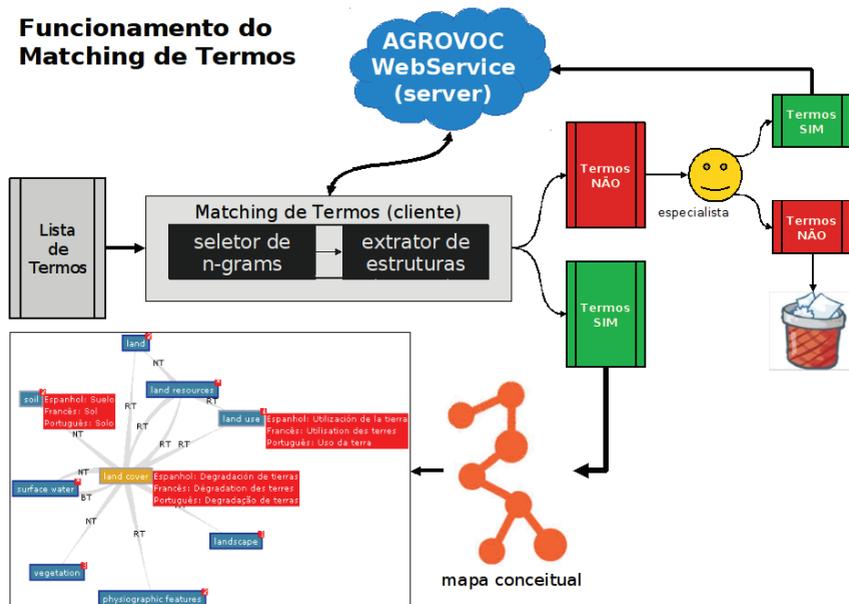
A interface foi configurada para recuperar termos em inglês (língua padrão do Agrovoc) e as respectivas variantes em determinadas línguas de origem latina (português europeu, espanhol, francês e italiano), permitindo que a comparação apoie a escolha das melhores opções em português brasileiro que podem ser incluídas no sistema. Conforme pode ser visto na Figura 1, o Etecam permite a realização das seguintes tarefas: dada uma lista de termos como entrada, a ferramenta perscruta no Agrovoc se aqueles termos já existem. Se sim, os termos são extraídos com as suas relações hierárquicas e associativas, bem como com os equivalentes em espanhol, francês, italiano e português europeu. Os termos que não existem no Agrovoc são enviados ao especialista de domínio que faz a seleção e, no caso de se tratar de um termo da Agricultura, o elege para a inclusão futura no Agrovoc, conforme mostra a Figura 2.

---

<sup>3</sup> Disponível em: <[http://www.embrapa.br/a\\_embrapa/missao](http://www.embrapa.br/a_embrapa/missao)>. Acesso em: 13 set. 2013.

<sup>4</sup> Disponível em: <<http://www.etermos.cnptia.embrapa.br/agrovocsearch/#>>. Acesso em 13 set. 2013.

### Funcionamento do Matching de Termos



mapa conceitual com relações semânticas e equivalências

Figura 1. Esquema de funcionamento do Matching de termos Etecam.

#### Consulta de Termos em Ordem Alfabética

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Inglês	Português (PT)	Espanhol (ES)	Francês (FR)	Italiano (IT)	Português (PT-Br)
(...)					
Cereal banks	Banco cerealífero	Banco cerealista	Banque de céréales	Banche di cereali	Banco de cereais
Cereal byproducts	Subproduto de cereais	Subproductos de cereales	Sous-produit de céréales	Sottoprodotti dei cereali	Subproduto de cereais, subproduto do cereal
Cereal crops	Planta cerealífera	Cultivos de cereales	Plante céréalière	Culture cerealicole	Cultivo de cereais, cerealcultura, cultura cerealífero, cultura de cereais, cultivo cerealífero
Cereal flours	Farinha de cereal	Harinas de cereales	Farine de céréale	Farine di cereali	Farinha de cereais, farinha de cereal
Cereal germs	Gérmen de cereal	Germenos de cereales	Germe de céréale	Germi di cereali	Gérmen de cereais, gérmen de cereal
Cereal products	Produto à base de cereal	Productos derivados de cereales	Produit céréalièr	Prodotti derivati dai cereali	Produto à base de cereais, produto à base de cereal, produto derivado de cereais, produto derivado do cereal

(...)

Figura 2. Inclusão de termos no Agrovoc Search.

Atualmente, os resultados advindos da aplicação desta metodologia aos subdomínios da Intensificação Agropecuária e Agroecologia têm-se mostrado bastante positivos; pois além das propostas dos mapas conceituais extraídos automaticamente, foram identificados e incluídos 1213 novos termos e/ou conceitos no Agrovoc na variedade brasileira do português.

Tais resultados mostram que a evolução desse trabalho tem viabilizado a execução de etapas básicas de um processo mais amplo que visa atender as seguintes demandas em relação às terminologias agropecuárias: 1) o atendimento às exigências do AO90; 2) a harmonização terminológica em língua portuguesa no domínio da Agricultura (pareando as variedades europeia e brasileira); e, principalmente, 3) a sistematização da terminologia brasileira para a Agricultura e domínios afins. Além disso, no nível mais operacional, possibilita que as ferramentas de automatização de algumas atividades desse processo sejam utilizadas em outras tarefas terminológicas.

## Literatura recomendada

AGROVOC Thesaurus. **Agricultural Information Management Standards (AIMS) – Food and Agriculture Organization of the United Nations (FAO)**. Disponível em: <<http://aims.fao.org/standards/agrovoc/functionalities/search>>. Acesso em: 14 out. 2013.

ALMEIDA, G. M. de B. Terminologia: o que é e como se faz. In: GONÇALVES, A. V.; GÓIS, M. L. de S. (org.). **Ciências da linguagem: o fazer científico?** Campinas: Mercado de Letras, 2012, v. 1, p. 197-230.

ALMEIDA, G. M. de B.; FERREIRA, J. P.; CORREIA, M.; OLIVEIRA, G. M. de. Vocabulário ortográfico comum (VOC): constituição de uma base lexical para a língua portuguesa. **Estudos Linguísticos**. São Paulo, 2013, v. 2, p. 204-215.

BRASIL. Ministério da Educação e Cultura. **Acordo ortográfico da Língua Portuguesa de 1990**. 1990. Disponível em: <<http://portal.mec.gov.br/arquivos/pdf/acordoortografico.pdf>>. Acesso em: 16 out. 2013.