

A new semantic resource responding to the principles of open science: the meat thesaurus for dialogue between sector actors

Moise Kombolo ¹, Sophie Aubin ², Jeremy Yon ², Francois Landrieu ³, Brigitte Richon ³, Jean-Francois Hocquette ¹

¹ INRAE, UMR1213, Recherches sur les Herbivores, ² INRAE, DipSO, 42 rue Georges Morel, 49070 Beaucozoué, France, ³ Académie de la Viande, 207 rue de Bercy, 75012 Paris, France

The Open Science Policy aims at developing new approaches for the scientific process. This is based on largely diffusing knowledge by using digital technologies and new collaborative tools. To achieve this, it is important to make the results of scientific re- search accessible in a simple and understandable way. Terminological resources like thesauri play a crucial role to succeed in this as they facilitate the collaboration between disciplinary and linguistic communities, the discovery of new information and data from various sources and the development of tools based on knowledge of the field. The purpose of the present work was to transform the Meat Dictionary published in 2012 by the French Meat Academy as a book into a machine actionable and freely accessible terminological resource, the bilingual French-English Meat Thesaurus. A thesaurus is a controlled and structured vocabulary in which concepts are represented by terms, organized in a hierarchy so that relationships between concepts are made explicit, and preferred terms are accompanied by entries for synonyms or nearly synonym terms. Textual definitions are provided in French and English for each concept. Concretely, we had to transform the original dictionary in Word format (.docx) into a thesaurus in RDF format using the SKOS (Simple Knowledge Organization System) standard model. We used a text editor with advanced functionalities and Open Refine to parse, transform, clean and correct the data. Finally, the data were then extracted to build a correspondence scheme between a tabulated format and a representation in SKOS. Once the format transformation operated, we imported the data into a semantic resource editor. This work was carried out by experts in semantic web, meat biology and meat vocabulary. The Meat Thesaurus contains 1567 concepts organised in 12 main categories and 145 sub-categories :1) “breed” includes the different breeds of cattle, goats, horses, sheep and pigs; 2) “butchery animals” includes animals that are slaughtered for meat consumption;

3)“culinary preparation of meat” includes concepts that describe meat-based preparations; 4) “cutting” describes all cuts for butchery animals (poultry, large cattle, lamb, equines, pork, rabbit and veal); 5) “game” includes all animals that are hunted for their meat and are not animals for slaughter (e.g. pheasants, kangaroos, hares, wild ducks and wood pigeons); 6) “health policy” covers different second-level concepts (including HACCP, diseases, veterinary inspection and traceability); 7) “livestock production” includes both generic livestock farming practices (e.g. related to “animal welfare”) and practices specific to a type of livestock farming; 8) “meat” is the core subject in this thesaurus, which covers the different types of meat (such as beef, game, poultry, sheep and pigs) and also provides various definitions and other notions (such as packaging, preservation, processing of muscle into meat, meat industry and meat quality); 9) “meat market” describes the commercial exchanges within the meat sector; 10) “meat professions” includes all types of occupations in the sector as well as training and distinctions; 11) “professional organisations” presents the different organisations divided into four categories (commercialisation-industry, distribution, interprofessional, production) and finally, 12) “slaughtering” contains the actions carried out in the slaughterhouse (stunning, killing, throat cutting, dehairing, evis- ceration, etc.), the tools used, ritual slaughter (halal and kosher), and concepts relating to the carcass (carcass yield, meat yield, at- las, fifth quarter, trapping, etc.) and carcass grading. This thematic organisation offers a navigation path that facilitates its discovery and use. This thesaurus can be used to index articles, journals and datasets, thus facilitating consultation. It can also be used to facilitate interoperability of indexed datasets and provide contextual definitions for building ontologies, that is, formal descriptions of knowledge for reasoning on data. The thesaurus can also be useful to enrich other vocabularies with new knowledge, such as French specificities in terms of meat cuts or definitions. In conclusion, the Meat Thesaurus is a resource that covers a wide range of concepts in the meat sector. Alone, or in combination with other existing and freely available resources, it will also make it possible to deepen a particular aspect of meat, such as its quality and will also enable reasoning for computers for a more automatic use. It is made available and free for use on Agroportal : <http://agroportal.lirmm.fr/ontologies/MEAT-T/>. More generally, the Meat Thesaurus agrees with the Open Science policy and FAIR principles promoted for instance in EU-funded research projects.

Key words: Meat, Thesaurus, Open science