Computers and Electronics in Agriculture xxx (xxxx) xxx

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

AGROVOC: The linked data concept hub for food and agriculture

Imma Subirats-Coll [a], Kristin Kolshus [a], Andrea Turbati [a], Armando Stellato [b], Esther Mietzsch [c], Daniel Martini [c], Marcia Zeng [d]

[a] Food and Agriculture Organization of the United Nations, Italy
[b] University of Rome Tor Vergata, Italy
[c] Association for Technology and Structures in Agriculture (KTBL), Germany
[d] Kent State University, United States

ARTICLE INFO

Keywords:
Agricultural Sciences
Research Data
Linked Open Data
Semantic Web
Interoperability
Data

ABSTRACT

Newly acquired, aggregated and shared data are essential for innovation in food and agriculture to improve the discoverability of research. Since the early 1980's, the Food and Agriculture Organization of the United Nations (FAO) has coordinated AGROVOC, a valuable tool for data to be classified homogeneously, facilitating interoperability and reuse. AGROVOC is a multilingual and controlled vocabulary designed to cover concepts and terminology under FAO's areas of interest. It is the largest Linked Open Data set about agriculture available for public use and its highest impact is through facilitating the access and visibility of data across domains and languages. This chapter has the aim of describing the current status of one of the most popular thesaurus in all FAO's areas of interest, and how it has become the Linked Data Concept Hub for food and agriculture, through new procedures put in place.

1. The demand for globally unique, unambiguous identification in agriculture

Production of food is a knowledge and information intensive process. In the past, farmers could manage most aspects of production by acting intuitively based on observations, experience and traditional knowledge handed on from one generation to the next. Over time, supply chains have become more complex. Available decision options in terms of production and marketing have increased. Additionally, the division of labour and duties in a network of stakeholders including farmers, suppliers, subcontractors, food processors and distributors has become a common principle in securing food supply to the population. In parallel, methods of capturing and distributing data and information have improved: Farm management as well as agricultural research were initially based on recording on paper in e. g. record card systems and laboratory notebooks, respectively. Printed publications on research findings and printed tabular data on variety field trials and archiving in libraries were a standard. Now, we see more and more the move towards digital systems using sensors for data capture, databases for storage and internet protocols and the web as a means of distribution.

These gradual changes in organizational setup of food supply as well as technical developments led to an increasing demand for data exchange, processing and information retrieval. Before it became a standard that computers were connected through TCP/IP, systems were mostly information islands. Nevertheless, data exchange was already conducted through physical storage media like tapes, floppy discs and optical storage media that were handed. To enable efficient data exchange between different databases, people invented data dictionaries and coding systems assigning simple alphanumeric codes for example to products, varieties, breeds or crops. That approach is still in place in a number of legacy systems used in agriculture like in the ISOBUS/ISO11783 data dictionary [ISOBUS, 2020] used in exchange between farm management information systems and agricultural machinery, the ICAR breed code list used for animal tracking purposes [ICAR] or the EPPO codes of crops used for plant protection applications [EPPO Codes, 2020]. Alas, working out an agreement on meaning and structure of these kinds of codes is a cumbersome process of committee discussions and providing textual documentation that differs for each of these coding systems as long as no formalized semantics following standards are used. As a result, implementing communication processes between systems requires a lot of effort as developers have to get acquainted with the coding systems and have to read a lot of documentation - especially if the use case to be tackled requires usage of data from different sources encoded in different coding systems.

To make matters worse, different coding systems were invented for the same "things". For example, there are a number of coding systems in place, that in some way refer to the concept of "crops" - each one defined for a different use cases:

- The EPPO code system

https://doi.org/10.1016/j.compag.2020.105965
Received 13 May 2020; Received in revised form 3 September 2020; Accepted 19 December 2020

0168-1699/© 2020 The Food and Agriculture Organization of the United Nations. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND IGO license (http://creativecommons.org/licenses/by-nc-nd/3.0/igo/).

- The USDA crop codes [USDACC]
- The Coding System used for the Integrated Administration and Control System (IACS) of the European Union [IACS]
- Code list used by different variety offices ranging from national to international level [UPOV]
- numerous others of minor importance…

Thus although that approach helped in implementing digital communications at a bilateral level or in systems and exchange processes with a very limited number of different agents, it did not scale to the networked world and data integration requirements of today and even prevented achieving interoperability.

2. The role of Linked data and semantic web vocabulary standards

The problem described is not specific to the agricultural domain and thus was recognized by web experts as well. It was part of the motivation for Tim Berners-Lee to postulate the principles of the web of data and linked data (Berners-Lee, 2006) and of the world wide web consortium to issue implementation specifications like RDF.

One of the core principles of the web of data is to push the use of URIs [Berners-Lee, 2020] as identifiers for anything. That way, varying rules for the formation of alphanumeric identifiers that might lead to overlaps and clashes between different coding systems are replaced by clear and standardized syntactical and management rules that guarantee global uniqueness as long as only the standards are followed - and that without the need for any effortful agreements between communities.

In practice all of the legacy coding systems existing in agriculture can be migrated to the use of URIs. In most cases, it will only involve the declaration of a prefix according to the syntax described in RFC3986 to form either a Uniform Resource Locator (URL) or a Uniform Resource Name (URN). For data integrators, that solves the problem of having to deal with different syntaxes for identifiers in controlled vocabularies and learning their formation rules. However it does not solve the problem of different controlled vocabularies for the same “things” still existing. Having foreseen that it is practically impossible to enforce unification towards a single, world-wide master controlled vocabulary, editors of the semantic web specifications introduced mapping properties in the recommendations:

- The web ontology language (OWL Working Group, 2012) provides the properties owl:sameAs, owl:equivalentClass, owl:equivalentProperty for instances, classes and properties, respectively
- The Simple Knowledge Organisation System (Miles and Bechhofer, 2009) provides the properties skos:mappingRelation, skos:closeMatch, skos:exactMatch, skos:broadMatch, skos:narrowMatch and skos:relatedMatch

Using these properties, it is at least possible to specify mappings between different controlled vocabularies in a unified and standardized way that formalizes these mappings and unlocks them towards being automatically interpreted by machines without human intervention or effort. It thus becomes feasible to systematically scan vocabularies, identify duplicates and merge them and the data sets using them using algorithms to automate that process.

Since controlled vocabularies are used to link data on the web and act as reference systems, it is nevertheless especially important to avoid as much duplication as possible already now. However, reusing existing URIs from existing vocabularies is not always an option, as each system or community has its own needs. This leads each community and even each institution to maintain their own silos, which means duplication of effort and costs.

From the technical point of view, although Linked Data is based on a distributed architecture which deals with redundancy and handles it with links, the proliferation of vocabularies’ re-defining the same concepts has become a serious problem for data curators. The choices are limited: either they link to the same thing already declared through several URIs defining the same concept or they choose one URI from one single vocabulary, perhaps the most authoritative, if there is one, relying on the fact that that URI will link to the others. This may work if vocabularies are indeed all linked, but building crosswalks and federated searches looking up several vocabularies is not easy. Besides, links themselves should be maintained and updated as the target vocabularies evolve.

From a management point of view, it is a considerable effort to select and define terms, and to translate and organize them. This requires consultations, users’ feedback and input from language experts. For different vocabularies, which cover a certain percentage of common concepts, the same effort is repeated, potentially involving the same experts. However, these results may differ and lead to different definitions, translations and organization of knowledge. In some cases these differences are essential and deserve different vocabularies, but in other cases communities may develop similar vocabularies since they work in silos and don’t reuse what already exists.

In the case of AGROVOC, when discussions started on reuse and integration, the thesaurus was already very broad and very authoritative, especially considering that it was managed by the Food and Agriculture Organization of the United Nations (FAO), the most authoritative organization for agricultural knowledge. This chapter summarizes the last procedures developed in the context of AGROVOC and its new role in linking data on the web.

3. AGROVOC

With the objective to increase accessibility and visibility of research products in its member countries, FAO has been promoting the exchange of scientific and technical information related to all aspects of agriculture since the early 1970’s. Moving from paper to digital and from processing to partnerships, FAO has established a series of programmes to support these efforts and make knowledge, information and research data on agriculture and related sciences available, accessible and usable.

First published in the early 1980′s by FAO to describe documents and other information resources in a controlled language for indexing and searching, AGROVOC has moved from print catalogues and databases to semantic web technologies. The multilingual AGROVOC thesaurus is now online and linked to other multilingual knowledge organization systems, building bridges between datasets. The advantage of having a thesaurus like AGROVOC published as a Linked Open Data set is that once vocabularies are linked, the resources they index are linked as well. AGROVOC is the most used thesaurus in agriculture and related sciences in the world, and is one of the most popular resources to index and tag information resources in low-income countries. For example, AGROVOC is the controlled vocabulary used to index AGRIS[1] records.

AGROVOC content is accessible both to humans and in machine-readable format. It can be searched manually for concepts or terms, browsed by hierarchy, downloaded as an RDF dataset, accessed via web services or searched through SPARQL queries, using a public SPARQL endpoint. AGROVOC concept URIs resolve to web resources delivering more detailed information like labels and relations to other terms for a certain concept(see Figs. 1–5).

AGROVOC provides a way to organize knowledge for subsequent data retrieval. It is a structured collection of concepts, terms, definitions and relationships. Concepts represent anything in food and agriculture, such as maize, hunger, aquaculture, value chains or forestry. These concepts are used to unambiguously identify resources, allowing standardized indexing processes, making searching more efficient. Each concept in AGROVOC also has terms used to express it in various languages, so called lexicalizations. Today, AGROVOC consists of + 38,100

[1] http://agris.fao.org/

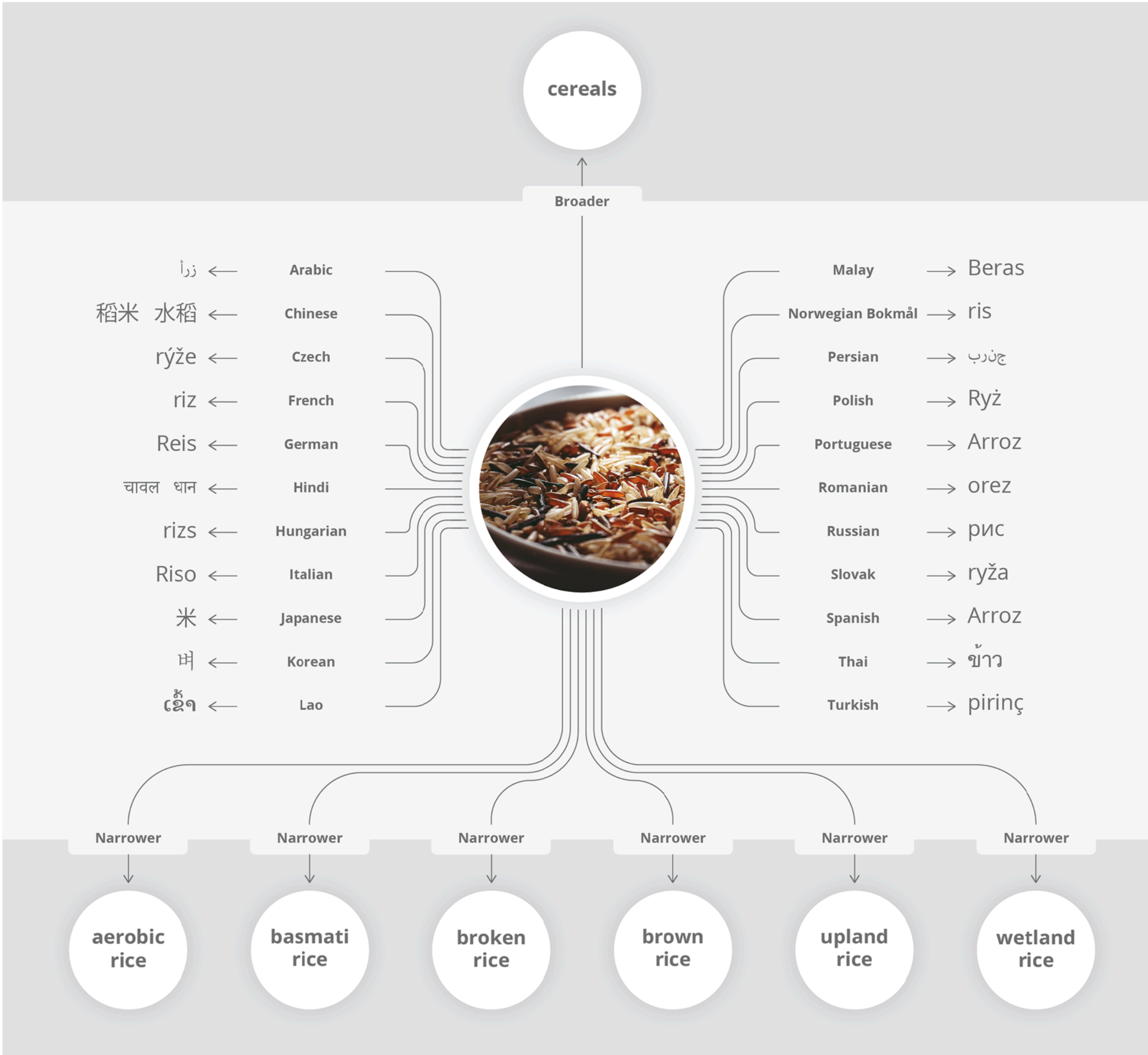

Fig. 1. AGROVOC Concept.

concepts and + 802,000 terms in up to 40 languages. AGROVOC is the largest thesaurus published as linked open data about food and agriculture available for public use.

As of April 2020, AGROVOC is available in Arabic, Burmese, Chinese, Czech, Danish, Dutch, Finnish, Georgian, German, Hindi, Hungarian, Italian, Japanese, Khmer, Korean, Lao, Malay, Norwegian Bokmål, Persian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Spanish, Swahili, Swedish, Telugu, Thai, Turkish, Ukrainian, and Vietnamese. Language coverage varies, The framework is in place for regional variants of Spanish and Portuguese, as well as Albanian, Belarusian, Bulgarian, Catalan, Croatian, Estonian, Gaelic, Greek, Indonesian, Latvian, Lithuanian, Northern Sami, Norwegian Nynorsk, and Slovenian, pending identification of national editors.

Copyright for AGROVOC content in FAO languages - English, French, Spanish, Arabic, Russian and Chinese - is with FAO, while content in other languages rests with the institutions that authored it. AGROVOC thesaurus content in English, Russian, French, Spanish, Arabic and Chinese is licensed under the international Creative Commons Attribution License (CC-BY IGO 3.0).

Since 2009, AGROVOC has been a SKOS thesaurus [OWL Working Group, 2012] and is Linked Open Data since 2018. It uses both hierarchical and non-hierarchical relations among concepts. Hierarchical relations among concepts are expressed by the predicates skos:broader and its inverse skos:narrower. Non-hierarchical ones vary on a wide range of semantic relations among concepts. In particular, AGROVOC uses the SKOS relation skos:related (corresponding to the classical thesaurus RT) together with a dedicated vocabulary of relations, called Agrontology, pertaining to the agricultural domain. Furthermore, AGROVOC, thanks to the adoption of the SKOS-XL (“SKOS eXtension for Labels”) vocabulary, provides descriptors for labels as “first class citizens’’, being thus reified as objects of the domain and able to be described in turn. This allows for the introduction of attributes for labels (mostly used for editorial purposes) and for the definition of relations among labels, thus purely addressing the lexical level. For instance, relations such as “has_acronym” univocally bind the labels “Food and Agriculture Organization” and “FAO”, with no concern for the organization itself or for other labels that it might have.

Agrontology is used to describe non-hierarchical relations between concepts: for example, “irrigation” is practice for “flooded rice” and influences ”soil salinization”, while “rain” causes “lodging”. “Lodging” causes “crop losses”, while “fins” are used in “locomotion”. These non-hierarchical relations express a notion of “relatedness” between concepts. When applied correctly, these additional relationships provide more context and richness to the dataset.

The highest impact of AGROVOC is through its role in facilitating the access and visibility of data across domains and languages. As

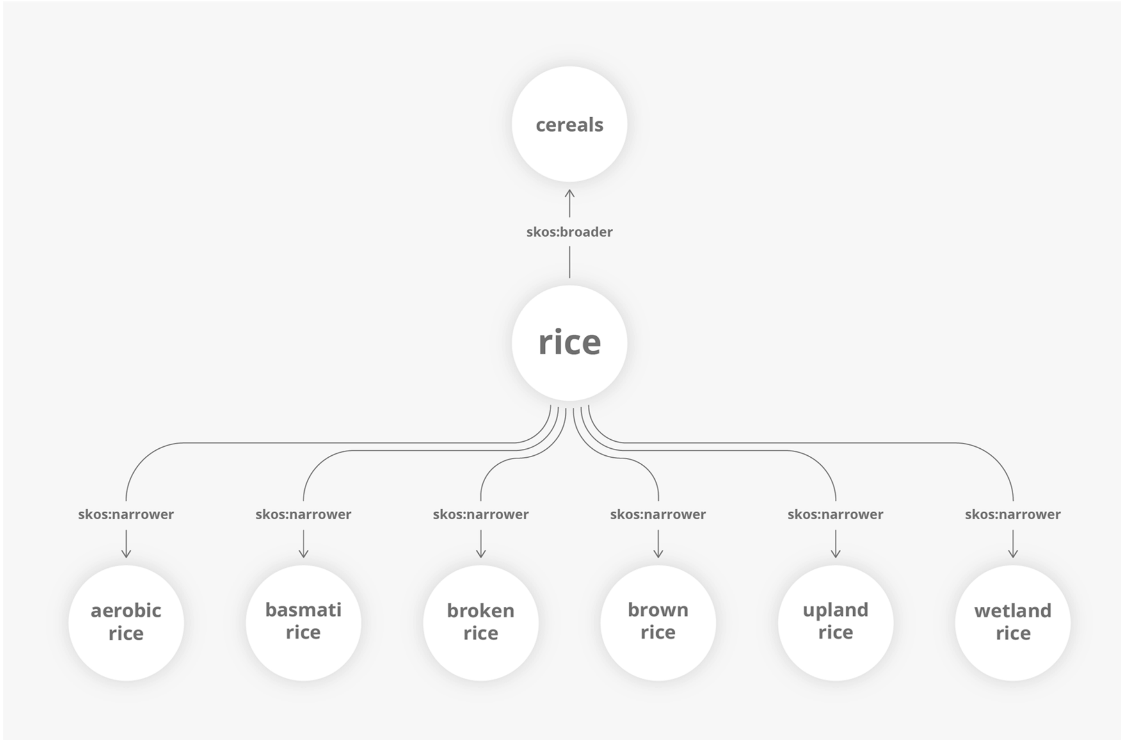

Fig. 2. AGROVOC Hierarchical Relationships.

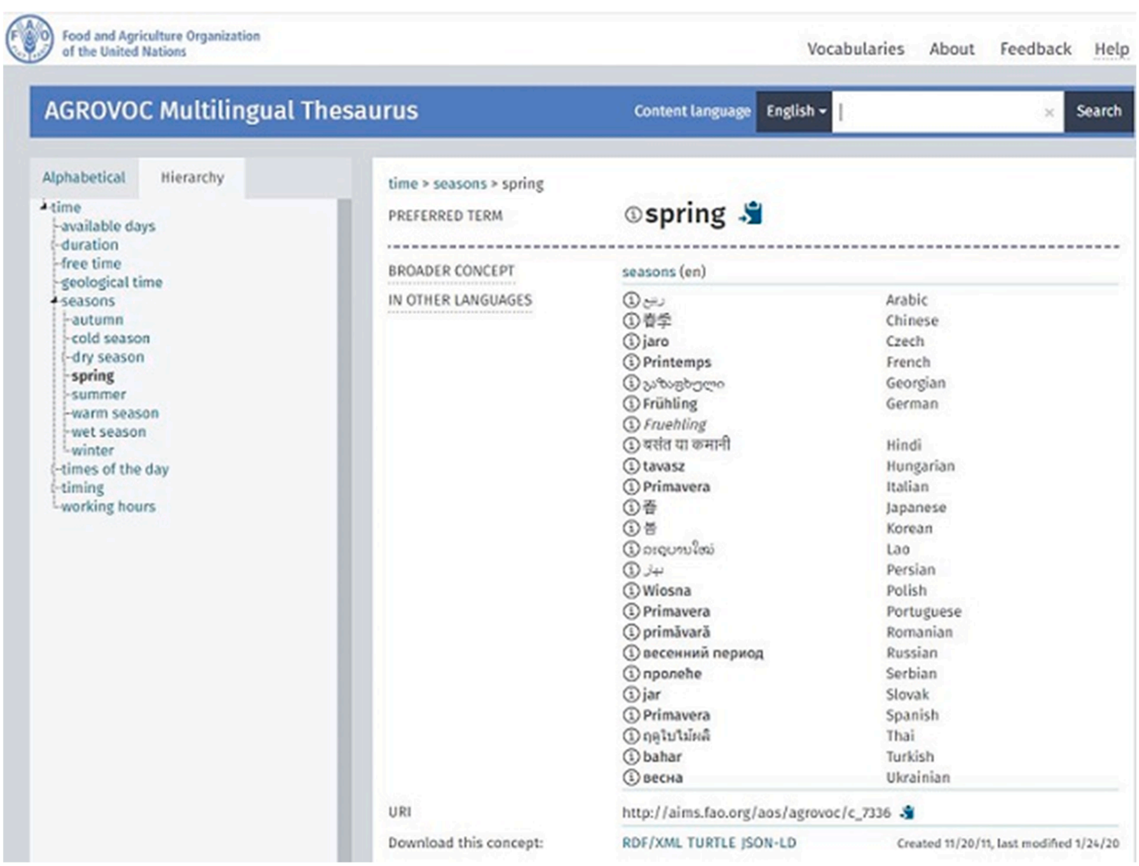

Fig. 3. AGROVOC Skosmos interface.

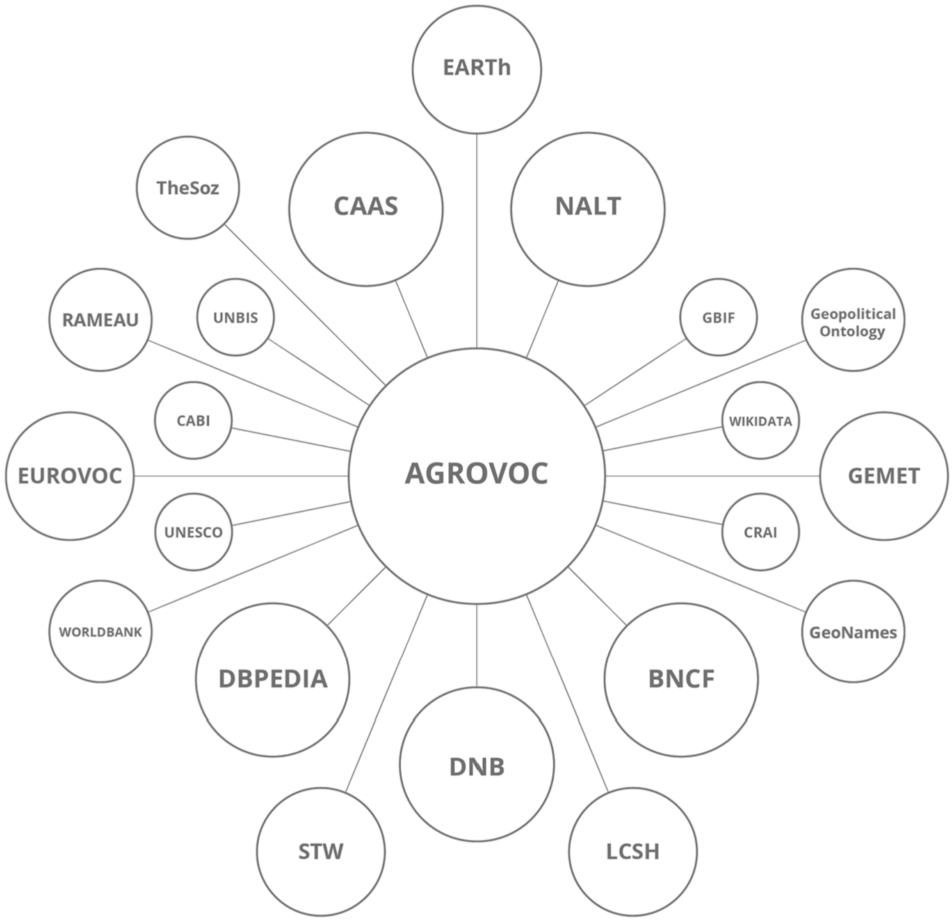

Fig. 4. AGROVOC Linked Open Data.

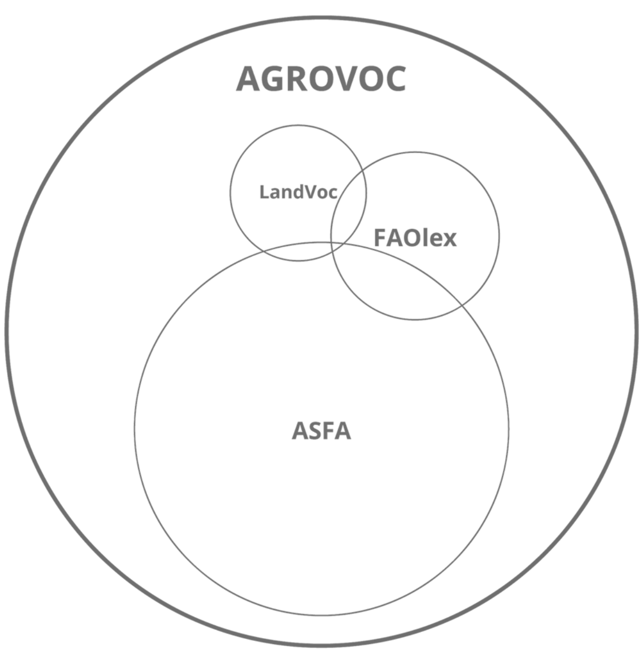

Fig. 5. AGROVOC and Schemes.

highlighted by Dean Allemang (Allemang, 2020), a Semantic Web expert: "Vocabularies like AGROVOC play a special role in the data ecosystem. They provide touch points where multiple data sets can refer to the same thing, and anyone who reuses those data sets knows, without reading through documentation or contacting the data set authors, that they refer to the same thing. They are key shared resources that allow other resources to interoperate"[2].

AGROVOC can be accessed in different ways[3] depending on the user's needs. AGROVOC editors use VocBench to edit AGROVOC. Human users mainly consult the SKOSMOS GUI[4] to search concepts and navigate the hierarchy. Legacy tools still interact with SOAP Web Services, while automatic tools consult SKOSMOS APIs, URI resolution or the SPARQL endpoint.

In 2019, AGROVOC had 10 million visits to AGROVOC concepts per year, 3.65 million access to AGROVOC search per year, and 500,000 accesses to AGROVOC Web services.

4. Collaborative and decentralized vocabulary maintenance

AGROVOC is a collaborative effort. The AGROVOC team at FAO bears responsibility for maintaining AGROVOC in the six languages of FAO (English, French, Spanish, Arabic, Chinese and Russian) and for coordinating all other contributions from the AGROVOC community of editors, which encompasses the maintainers of AGROVOC language

[2] Dean Allemang. *Sustainability in Data and Food*. Data Intelligence 2019 1:1, 43–57

[3] http://aims.fao.org/agrovoc/webservices

[4] http://agrovoc.uniroma2.it/agrovoc/agrovoc/en/

versions and communities of experts in specific domains. This work is essential, with national and international organizations volunteering to be responsible for the different language versions and subject domains. Their motivation is both to utilize AGROVOC in their own language in their national systems, and to increase visibility of research in their language to a global audience.

Contributing institutions include: Agroinstitut Nitra (Slovakia), Biblioteca Storica Nazionale dell'Agricoltura (Italy), BonaRes Project, Helmholtz Centre for Environmental Research (Germany), Central Scientific Agricultural Library (Russia), Chinese Academy of Agricultural Sciences (China), CIRAD (France). Embrapa (Brazil), Department of Training, Extension and Publications, Ministry of Food Agriculture and Livestock (Turkey), Gödöllo Agribusiness Centre (Hungary), ICARDA, Institute of Agricultural Economics and Information (Czechia), Iranian Fisheries Science Research Institute (Iran), Kasetsart University (Thailand), Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (Germany), Land Portal Foundation, Matica Srpska Library (Serbia), TECHINFORMI (Georgia), and the Republican Scientific Agricultural Library, State Agrarian University (Moldova).

The AGROVOC approach to collaborative and decentralized vocabulary maintenance has also been highlighted by NISO: "In the future, we can imagine a broadly distributed ecosystem for vocabulary creation, maintenance, and use based on a commonly agreed URL infrastructure, built to support distribution of terms to consumers based on their explicit preferences. The Food and Agriculture OrganizationFAO implements such a model for AGROVOC and it is instructive to review its features.[5]"

Since 2019, AGROVOC is expanding its coverage through the collaboration with communities of experts to include specialized domains who can benefit from the AGROVOC infrastructure: new vocabularies embedded in AGROVOC, enriching AGROVOC, while maintaining the possibility for separate entity, exports and display. This engagement with specialized communities is helping AGROVOC to enhance the quality of AGROVOC content, while the new vocabularies benefit from the AGROVOC community of editors, e.g. translations and quality control. This new AGROVOC paradigm can be useful for smaller specialized technical vocabularies to leverage semantic technologies and linked open data. This supports agricultural sciences data to become more FAIR (Findable, Accessible, Interoperable, and Reusable) as defined by the Go FAIR initiative[6].

5. Supporting collaborative editing: The VocBench platform

AGROVOC is edited through the web-based platform VocBench[7] (Stellato et al., 2020, 2017), a free and open-source advanced collaboration environment for creating and maintaining ontologies, KOSes (thesauri, code lists, authority resources, etc..) and lexicons. VocBench (or, more shortly, VB3) supports the collaborative development of these resources by completely embracing Semantic Web standards, such as OWL, SKOS (SKOS-XL) and Ontolex-Lemon, for content representation, metadata standards (e.g. VoID, LIME, DCAT and its AP variations for the different EU countries, ADMS) for content publication, dissemination and other vocabularies that assist content creation, validation and maintenance in general (e.g. SHACL for constraint validation, SPARQL protocol and query/update language, etc..).

Besides mere editing and browsing capabilities, VB3 features several advanced functionalities for supporting the publication workflow (e.g. history, validation and versioning), ETL processes (e.g. import from spreadsheets), alignment (manual, semi-automatic and automatic), user discussions (through collaboration tools such as JIRA), Integrity Constraint Validation and extension points for allowing third party developers to extend all of these functionalities with new implementations for different technologies and scenarios.

VB3 is funded by Action 1.1 of the ISA[2] Programme of the European Commission for "Interoperability solutions for public administrations, businesses and citizens"[8]. The action is managed by the Publications Office of the European Union[9] and developed by an industry-academy joint collaboration involving the University of Rome Tor Vergata and two IT companies: Infeurope and Lore Star

VB3 was released for the first time to the public on September 2017, under a BSD 3-clause license[10]. Since then, an average of 2–3 releases each year marks constant evolutions of the system characterized by the introduction of new features, improvements over the previous ones and, obviously, bug fixes.

The VocBench site[11] contains documentation, download links and other references.

5.1. AGROVOC and VocBench: A binary star

The stories of AGROVOC and VocBench are intimately connected as the former inspired and guided the development of the latter since its infancy. Everything started, in 2008, when FAO, as a partner of the NeOn project, decided to foster the development of a collaborative platform for managing AGROVOC. NeOn was a 4-year project, which started in 2006, involving 14 European partners, co-funded by the European Commission's Sixth Framework Programme under grant number IST-2005–027595. The aim of the project was to advance the state of the art in using ontologies for large-scale semantic applications in distributed organizations, with a particular focus on handling multiple networked ontologies, managed collaboratively, with a highly dynamic and constant evolution.

Collaboration, distributed organizations and constant evolution were all key aspects perfectly matching with the AGROVOC scenario. However, one important ingredient was missing: AGROVOC is a thesaurus and the project was going full throttle on OWL ontologies. Besides the simple but radical differences in the adopted models (which could be excused at the time, the release of SKOS specifications dates to 18 August 2009, three years after the start of the project) other key factors, mostly connected to the user experience and collaboration, led to the idea of developing a dedicated system.

5.2. The system-with-many-names

In 2008 and with the approval of the consortium, the "Neon-light" was initiated as a collaborative development for thesauri and authority lists. The first release of the system was completed in 2009. At the time, the system was internally rebranded as "AGROVOC Workbench" aka "AGROVOC Concept Server". The system was based on Protégé3 [Musen, 2015], exploiting its API for the data backend while featuring a completely independent middle-layer and frontend developed with the GWT technology[12]. Indeed, there were still many technological limitations with respect to the needs of an RDF platform (Protégé 3 partially supported RDF thanks to an extension for OWL) and much of its code was strictly tailored to AGROVOC. Later on, the rising interest in the platform from other FAO departments and other organizations motivated its reengineering into a more general thesauri management system.

Sealing its firm direction as an editing system for thesauri, authority lists and, more in general, vocabularies, the latest - and last - name

[5] https://groups.niso.org/apps/group_public/download.php/18410/NISO_TR-06-2017_Issues_in_Vocabulary_Management.pdf
[6] https://go-fair.org
[7] http://vocbench.uniroma2.it/
[8] https://ec.europa.eu/isa2/
[9] https://publications.europa.eu/
[10] https://opensource.org/licenses/BSD-3-Clause
[11] http://vocbench.uniroma2.it/
[12] http://www.gwtproject.org/

"VocBench" was thus, eventually, coined.

5.3. Towards VocBench 3

The realization of this new incarnation of the system happened in the context of an existing collaboration between FAO and the ART group of the University of Tor Vergata[13] in Rome. With the objective of meeting the request of potential stakeholders of VocBench, a quick roadmap (6 months of development) was laid down foreseeing the merge of the original VocBench with ART's service-based RDF Management Platform Semantic Turkey [Pazienz et al., 2012].

The result of this collaboration, VocBench 2 [Stellato et al., 2015] had been rethought as a fully-fledged collaborative platform for thesaurus management, freely available and open-sourced, offering native RDF support for SKOS and SKOS-XL knowledge organization systems (Hodge, 2020), while retaining from its original version the focus on multilingualism, collaboration and a structured content validation & publication workflow.

VocBench 2 has been the first real "community" release of VocBench. Many important organizations, who were migrating or planning to migrate to VocBench, jumped on board such as: the Italian Senate of the Republic, for the maintenance of the Teseo thesaurus, the Publications Office of the European Union, for the maintenance of the EuroVoc vocabulary, the "Institut national de la recherche agronomique" (INRA) in France, some universities such as Columbia University (GEDA thesaurus) and Harvard (Unified Astronomy Thesaurus) in the US, among others.

VocBench 2 was actively developed, until March 2017, and was supported for the following two years. However, in order to support the larger audience gathered around VocBench, and the many different needs that it raised, in 2016, the Publications Office of the EU, supported by funding by the ISA2 programme of the EU, took over the management of the project: VocBench 3 was born. Today, VocBench 3 is actively developed by the same team as VocBench 2, in collaboration with the company infeurope[14] under management of the Publications Office of the EU. The range of covered knowledge models was extended in VB3 to cover also OWL ontologies, Ontolex-Lemon lexicons, EDOAL alignments and RDF datasets in general. Since November 2016, VocBench has been actively developed, foreseeing a minimum of two releases each year and introducing new features, with improvements and bug fixes.

5.4. AGROVOC and VocBench 3: A fruitful symbiosis

Even after the advent of VocBench 3 and the managerial move to the European Commission, FAO is still one of the main stakeholders of the project. With its non-trivial size, the presence of different concept schemes owned by different organizations co-working on the same dataset, its large number of supported languages and wide community distributed across the globe, AGROVOC represents one of the most important use-cases challenging VocBench under many profiles, among which: performances, usability, agility and richness of features.

Conversely, the highly dynamic evolution of VocBench and its proactive development provides a quick response to the AGROVOC demand for new features and new scenarios, mutually affecting each other's evolution in a positive way.

6. New AGROVOC Paradigm: A Linked data concept hub for food and agriculture

The scope of AGROVOC is vast: agriculture, fisheries, forestry, and environment. Managing the thesaurus content in specialized areas requires deep subject matter expertise; yet the AGROVOC core team is very small. Content curation depends on collaboration with an international network of institutions, traditionally curating a language. At the same time, there are a number of technical communities of practice which may have a controlled vocabulary in their area of research, but lack modern infrastructure to share this as Linked Data. This means that they are missing the opportunity to leverage the power of semantic technologies to make their research more accessible and visible. By inviting expert communities to add their controlled vocabularies and thesauri to the AGROVOC framework, AGROVOC benefits from this new subject matter expertise. At the same time, these collections benefit from AGROVOC technical infrastructure and editorial network, while also becoming part of the AGROVOC Linked Open data structure.

This opens for some interesting collaborations with specialized communities that can benefit from the AGROVOC infrastructure: embedding their vocabularies within AGROVOC, enriching AGROVOC, while maintaining the possibility for separate identities, exports and display. Expert communities can now curate a topic within AGROVOC. In practical terms, every single concept used by any of the schemes is part of the AGROVOC Thesaurus and has AGROVOC URIs.

Since 2019 and with the collaboration of the Artificial Intelligence (ART) Research Group at Tor Vergata University (Italy), the management of specialized concept schemes is possible within AGROVOC. VocBench 3 supports the use of hierarchical relation properties that are specific to a scheme. *However, w*hen a concept is modified (e.g. adding or changing a translation) in a given scheme, the data is not only edited for that scheme, but associated to a concept and such update will now be seen in all schemes which include this concept. That way, any such scheme can benefit from a common "concept pool" within AGROVOC.

Currently, AGROVOC includes three new schemes:

- ***Land Governance represented by the LandVoc, coordinated by the Land Portal Foundation.*** LandVoc is a controlled vocabulary covering any concepts related to land governance. It is the result of an ongoing process in which the land terms in AGROVOC have been enriched (using standards such as LADM or the Multilingual Land Tenure Thesaurus, but also local partner inputs and translations). LandVoc currently consists of 290 concepts and is available in 5 languages (English, French, Spanish, Swahili and Portuguese) and is continuously translated into more languages.
- ***Aquatic Sciences and Fisheries represented by ASFA, coordinated by the ASFA secretariat at FAO.*** Founded in 1971, ASFA is an international co-operative of fisheries and aquaculture organizations that work together to record and disseminate the world's literature on aquatic sciences and fisheries. Over 60 partners make up the ASFA partnership, all of whom compile bibliographic records which are available to search on the ASFA database. The overall objective of ASFA is to disseminate information on aquatic sciences and fisheries to the world community, bearing in mind the special interest of low-income countries in this field. The ASFA Vocabulary is an indexing tool that contains the subject descriptors used to index the records which are contained in the ASFA database.
- ***Legislative and Policy concepts in the FAO's areas of interest represented by FAOLEX, coordinated by FAOLEX.*** FAO has been committed to the collection and dissemination of legal, governance and policy information for more than 70 years. Its core mandate has always included the collection, analysis, interpretation and dissemination of information relating to nutrition, food and agriculture. Administered by the Development Law Service (LEGN) of the FAO Legal Office, FAOLEX complements FAO's core function of advising its Members on legal and institutional means to promote and regulate national development and international cooperation in the food and agriculture sector.

The engagement with specialized communities is helping AGROVOC to enhance the quality of its content, while the new schemes benefit from the AGROVOC community of editors, e.g. translations and quality

[13] http://art.uniroma2.it

[14] https://www.infeurope.lu/

control. The main consideration when dealing with multi schemes in AGROVOC is that each concept always belongs to the main scheme (AGROVOC), and it must be in a specific place within the AGROVOC hierarchy. However, each scheme might want to use different hierarchy models.

AGROVOC offers a flexible environment where schemes can define additional hierarchical relationships among concepts, independent of AGROVOC hierarchy. This means that a controlled vocabulary can be viewed flexibly and edited with its customized relations, or exported with a generic SKOS hierarchy of broader and narrower relations, without changing the hierarchy of AGROVOC itself.

When a concept URI (URI resolution) is being automatically resolved, the display is related to the AGROVOC concept. Therefore the displayed hierarchy will be from the AGROVOC main scheme. The rest of scheme specific hierarchical relations will not be visible. Available upon request, however each scheme can be viewed flexibly and separately, edited with its customized relations, or exported with a generic SKOS.

Recently, since the reuse of data is unanimously recognized as a big driver for innovation and the way data is shared is key to its reuse, there has been new interest around a definition of a more formal and more coordinated framework that could cater more for data-intensive research and sharing data across the data value chain. In 2014, the need to better define the 'rules' for a more effective sharing of data led a group of representatives of different stakeholders – academia, industry, publishers, funding agencies – to meet in The Netherlands and discuss a "minimal set of community-agreed guiding principles and practices"[15]. What came out of these discussions was a set of principles called the FAIR principles: according to these principles data must be: Findable, Accessible, Interoperable and Reusable (FAIR).

The focus is on clear access rights rather than openness, in line with the need for more flexibility as well as more precision in data sharing. This is needed to facilitate also the sharing of data that may have some access restrictions, but that can still be reused under specific conditions. There is special attention for provenance and attribution and for persistence, in line with the fact that the FAIR principles have been agreed upon by a community that wants to work together and share data and needs a trusted environment with some basic rules.

7. Practical usage scenarios

7.1. Topic scope of AGROVOC

When introducing AGROVOC into implementation projects, a first step for integration involves checking to what extent the concept scopes of the application planned and the concepts available in AGROVOC overlap. As mentioned above, AGROVOC is a hierarchically arranged SKOS concept scheme using the skos:broader and skos:narrower relations to organize the tree. A closer look at the topmost concepts in the tree thus provides an overview of the scope:

- Activities: contains all kinds of activities that are conducted along the food supply chain, like "breeding", "feeding", "surveying", "cleaning", "transport" but also more high-level management activities like "accounting" and "planning", activities on nutritional topics like "weight reduction" and activities that are more loosely related to agriculture and food or rural areas like "cartography", "computer programming" or "recreation".
- Entities: entities are rather broadly defined as "something which is distinct and separate from something else." That includes narrower concepts like "agencies", "labels", "networks", "policies".
- Events: events in this context are outlined as something taking place at a certain point in time and involving the participation of people, so includes concepts like "exhibitions" and "training courses"
- Factors: in agricultural research and publications, the term "factors" is frequently used in a number of rather common word combinations and these common combinations are reflected in the narrower concepts to be found here, e.g. "abiotic factors", "biotic factors", "environmental factors" or "production factors".
- Features: "features" here relates to the feature concept from geosciences and genetics and contains only three direct narrower concepts: "genomic features", "physiographic features" and "soil morphological features".
- Groups: defined as "a number of individual items or people brought together." narrower concepts like "engineers", "librarians" but also societal groups like "consumers" and "interest groups" can be found here.
- Location: a location is a "a point or extent in space" and thus holds concepts like "climatic zones", "maritime zones", "protected areas" or "urban areas".
- Measure: While a measure can also denote an action taken, in this context it is clearly defined as something, that can be observed and involves a measurement: "Number or quantity that records a directly observable value or performance. All measures have a unit attached to them: inch, centimeter, dollar, liter, etc.". Examples of narrower concepts are: "altitude", "breeding value", "humidity", "price indices", "soil water potential" etc.
- Methods: methods describe ways of doing things, either in agricultural research or in production but also in everyday life. They are like recipes - and as a notable fact, "cooking methods" is a narrower concept of the methods top concept.Other examples include "autoclaving", "irrigation methods", "sampling", "statistical methods" or "survey methods". Methods differ from activities in that for the former usually a formal or standardized description is available and that they involve a number of steps.
- Objects: Objects in this context subsume man-made, touchable things like "equipment" and "furniture".
- Organisms: The Organisms tree is one of the largest subtrees in AGROVOC and contains the taxonomic trees of organisms relevant to agriculture under subconcepts like "Eukaryota" and "Prokaryotae" as well as common organism classes like "plants" and "animals", but also roles that an organism can hold like "hosts", "pests" or "predators". Also available are concepts for specifying organisms that live in a certain habitat like "aquatic organisms" or "soil organisms".
- Phenomena: In scientific usage, a phenomenon is any event that is observable, however common it might be, even if it requires the use of instrumentation to observe, record, or compile data concerning it. In natural sciences, a phenomenon is an observable happening or event.This tree contains concepts like "deficiencies", "economic phenomena", "hazards", "population dynamics" or "trends".
- Processes: A process is a set of interrelated or interacting activities which transforms inputs into outputs. Examples of processes' narrower concepts include: "anthropogenic changes", "biological processes", "evolution", "inhibition", "physiological processes" or "synthesis".
- Products: in the context of the AGROVOC, these are mostly confined to products and product classes originating from agricultural supply chains, like "animal products", "feeds", "foods" or "oil products". But also, raw materials or product properties play a role represented by concepts like "resins" and "forest products" or "biodegradable products" and "sustainable products" respectively.
- Properties: a property is a characteristic or quality that can be owned or posessed, serving to define or describe its possessor. This tree contains numerous narrower concepts of differing granularity, e. g. "age", "colour fastness", "periodicity", "soil properties", "toxicity" or "wind direction".

[15] Force11. Guiding Principles For Findable, Accessible, Interoperable And Reusable Data Publishing Version B1.0. https://www.force11.org/fairprinciples

- Resources: resources are things that are used during a production process or that are required to cover human needs in everyday life. Concepts like “economic resources”, “inputs” and “raw materials” would refer to the former category. The latter category is covered by more abstract resources like “cultural heritage” or “natural resources”.
- Site: sites contain narrower concepts that serve to describe locations and facilities that are setup by humans for a certain purpose like “hospitals”, “laboratories”, “meteorological stations”, “restaurants” and “timberyards”.
- Stages: The stages top concepts has only two narrower concepts: “developmental stages” and “life cycle”. The former concept however is highly branched, containing all kinds of plant and animal development stages like “embryo stage”, “reproductive stage” etc.
- State: any condition in which a physical substance or organism can be in. Some narrower concepts are: “anoxia”, “colloidal state”, “employment”, “physical states”, “sleep” or “welfare”.
- Strategies: strategies describe acting options and include communication, rural development and training strategies as well as “approaches”.
- Subjects: being disciplines of study or topics relevant to agriculture and nutrition in this context, it includes e. g. “cartography”, “humanities” and “sciences”.
- Substances: substances is a rather broad subtree providing hierarchies for chemical substances according to physical properties like “ceramics”, “explosives”, “oils” or “solutes” but also according to their role or function like “attractants”, “culture media”, “drugs” or “soil amendments”, their source or place of origin like “exudates”, “filter cakes” or “sediment”.
- Systems: The systems top concept contains a wide range of concepts for systems of human action, interaction and thought (“economic systems”, “political systems”, “value systems”), production and supply (“distribution systems”, “drinking water systems”, “agroforestry systems”), technological systems (“information systems”, “photovoltaic systems”, “surveillance systems”) as well as systematic and organizational approaches from science (“knowledge organization system”, “terminology”).
- Technology: subsumes all the concepts for technological developments and inventions that are applied in modern agricultural and food systems: “biotechnology”, “food technology”, “information and communication technologies”, “seed technology”, “wood technology” and so on.
- Time: contains concepts that describe timespans with a certain function - like e. g. “free time”, “seasons”, “times of the day”, “working hours” and timestamps relevant to agricultural production - mostly aggregated in the “timing” concept.

AGROVOC is currently the thesaurus with the broadest coverage of concepts in agriculture and food. A more detailed overview can be gained by navigating the hierarchy using the SKOSMOS frontend. Concepts missing for certain use cases can be added through contacting the editorial community provided that the editorial guidelines allow their addition. It is obvious that there are corner cases for which it is not easy to pinpoint their place in the hierarchy. For example, one might argue that a process is a series of steps that might also be defined as a method or activities might have structure enough to justify an assignment to the methods tree. Also, objects might be products at the same time and measures are always also properties etc. Such cases can and should be discussed within the editorial community.

7.2. Text annotation and indexing

With more and more text and data produced, findability of resources like publications, articles, data sets and services becomes a crucial factor for targeted and efficient information retrieval. In the past, libraries served as the key interfaces to such resources. They developed topic classification and standard keyword systems - basically sorts of controlled vocabularies - to enable efficient search in their catalogues. One of the major tasks of librarians was and still is to classify publications accordingly and assign the proper keywords. Knowledge organization standards like SKOS are based on the experiences gained with these kinds of controlled vocabularies and library systems. Text annotation systems and tools aim at facilitating the cumbersome identification of key topics and entities in a text necessary for proper classification. That way, it can provide functionalities that the simpler but still more common full-text indexing and search cannot:

The latter treats most of the words in a text alike. Bare full text indexing can thus support a relevant judgement of a text only based on the terms the user entered for search and statistical measures of their occurence. Quality of the results will thus depend largely on the search terms used. Text annotation in turn can make use of formalized controlled vocabularies and leverage the semantics encoded within them to derive for example broader or related concepts from keywords contained in a text or to also provide results for synonyms of the search terms entered or to deliver navigational aids in a text corpus through a hierarchical concept browsing interface.

A system leveraging indexing using AGROVOC is the FAO AGRIS system. AGRIS is a global, multilingual bibliographic database that connects users directly to a rich collection of research and worldwide technical information on food and agriculture containing more than **12 million bibliographic records produced by more than 500 data providers including research centers, academic institutions, publishers, governmental bodies, development programmes, international and national organizations from 150 countries**. It also provides full text links to about 3 million of its records. It facilitates access to publications, journal articles, monographs. book chapters and grey literature - including unpublished science and technical reports, theses, dissertations and conference papers in the area of agriculture and related sciences. AGRIS is used by agricultural and research professionals worldwide in their everyday work. Records can be submitted by organisations in a number of standard bibliographic metadata formats like e. g. Crossref, DOAJ, Endnote, MARC21, MODS, Simple DC, and PubMed as long as they conform to the Meaningful Bibliographic Metadata (M2B) recommendations [Subirats et al., 2012]. Records are annotated using AGROVOC allowing e. g. for multilingual keyword search based on AGROVOC concepts.

Another recent text annotation project aims at enhancing findability and accessibility in the pest alert text database of German agricultural advisory services, enabling systematic querying, evaluation and interpretation using computer algorithms. These texts have been issued to farmers by regional advisors since years using faxes in early days and e-Mail and the web now and an archive is maintained by the Julius-Kühn-Institute. The messages contain information on pest risks on certain crops based on parameters like growth stages, weather conditions, experiences from past years and prognosis model results. Additionally, they give recommendations on whether treatments against certain organisms are indicated and which pesticides or mechanical and crop management measures can be applied. As such, a number of AGROVOC concepts like narrower concepts of organisms, activities, factors, locations and measures are used in these texts. Relying on a controlled vocabulary to annotate the texts would allow more targeted search like e. g. selecting all the records from past years dealing with a certain pest and crop within a regional boundary. This functionality could contribute to developing better treatment strategies for the future by identifying successful and unsuccessful measures under given certain influential factors in the past or for assessing long term proliferation of pests due to changing environmental or climatic conditions. The project has started in 2019 and will rely on a mix of manual and automatic annotation using AGROVOC as the core controlled vocabulary backbone. First results might be available by the end of 2020.

Ever since information technologies have been used to archive and process texts and documents, tools for manual keyword assignment and annotation of records have been developed in parallel. In recent years, also software for automated annotation has come up. A notable example is the annif tool for automated subject indexing and classification (Suominen, 2019). It has been developed by the National Library of Finland and can annotate a text based on a given controlled vocabulary based on the SKOS recommendation. Its source code has been made available under the Apache 2.0 open source license[16]. Another tool is Apache UIMA that provides a broader framework for text annotation also allowing recognition of parts of text based on syntactical patterns using for example regular expressions[17].

7.3. Metadata annotation of data sets

Increasing amounts of data generated by various entities ranging from research organisations through companies to individuals are being published on the web. Without principles of organizing and managing these massive amounts of data this is however comparable to a library with media and books of various kinds and in various languages lying around in a completely messy manner on an enormous amount of chaotically arranged shelves. Finding and extracting information literally becomes the search for the needle in the haystack. The FAIR principles elaborated within the context of the European Open Science Cloud [Wilkinson et al., 2016] aim at bringing order into this mess by introducing features to be factored into data offerings that allow for Findability, Accessibility, Interoperability and Reusability. One of the steps involved in the FAIRification process of data is the definition of metadata, which should make use of “vocabularies that follow FAIR principles” (principle I.2 in https://www.go-fair.org/fair-principles/). AGROVOC is such a vocabulary and as such can be used to annotate not only text resources but also data resources on the web.

Data repositories are making use of AGROVOC for that purpose. The BoNaRes project in Germany is dealing with soil as a natural resource and its data repository[18] is providing a number of soil research data sets. Many of these data sets have a strong linkage to agriculture in that research sites are subject to agricultural activity and use. The defined metadata schema [Gärtner et al., 2017] allows assignment of AGROVOC concepts as keywords to describe data sets by giving e. g. measures, factors or activities.

7.4. Ontology and RDF schema integration

Agriculture as a research domain is highly interlinked and overlapping with other sciences like biology, meteorology, geography, machinery construction and engineering, economics and so on. Given that characteristic, one can observe that data sets of value to agricultural research and practice are reflecting that situation by being distributed and heterogeneous. As a consequence, data integration in agriculture is suffering from numerous incompatible format specifications and coding systems and informal or undefined semantics. One has to acknowledge that following two basic principles is crucial to success in achieving interoperability in such a setting: 1. Use of globally unique identification as described above 2. Use of a data model that can accomodate any existing data sets and formats - essentially a graph-oriented data model. While the Resource Description Framework (Schreiber and Raimond, 2014) - upon which AGROVOC is also based - was initially drafted to assign rich metadata to web resources, it follows exactly these two principles and thus increasingly serves as a general model for data representation as well. People have built large data sets completely represented as RDF. Translations for data in relational databases are possible using tools like D2RQ[19] or db2triples[20]. JSON-LD allows semantic enrichment and transferability into the RDF model for simple JSON formats. Finally, it is usually feasible to write custom converters for other legacy formats as long as a syntactic parser is still available. That way, RDF is more and more taking the role of an interoperability catalyst. Data definitions and formats can be semantically formalized using RDF Schema or OWL ontologies in RDF triple format. One of the key questions in improving linkage of various resources is how one can make use of large controlled vocabularies and thesauri like the AGROVOC in that context.

7.5. Values spaces of Properties: Codelists reloaded

Remembering all that has been said about the demand for globally unique identification in agriculture, we can ask ourselves how we can leverage AGROVOC concepts as “data field value codelists” or enumerations in an RDF Schema based data format description. Value spaces on a property can be defined in RDF Schema using rdfs:range. Selecting and assigning subsets from AGROVOC as these value spaces is a simple and straightforward process:

As a first step, the respective property has to be defined:
myrdfs:crop a rdfs:Property.
To assign a value space we need a class to be used on that property:
myrdfs:Crop a rdfs:Class.
We then need to declare myrdfs:crop to use instances of myrdfs:Crop as a range:
myrdfs:crop rdfs:range myrdfs:Crop.
We select the subset of AGROVOC concepts that we want to use explicitly by stating that they are instances of the class myrs:Crop:
<http://aims.fao.org/aos/agrovoc/c_7951 > a myrdfs:Crop . # wheat
<http://aims.fao.org/aos/agrovoc/c_7221 > a myrdfs:Crop . # potatoes
etc.

The graph containing these schema statements could be automatically generated using e. g. a SPARQL CONSTRUCT query, thus introducing new concepts as they are added to certain hierarchy branches in AGROVOC and keeping a format always up to date. There may be more elaborate approaches avoiding having to use classes as individuals at the same time (skos:Concept is defined to be an instance of owl:Class in the SKOS specification) or using OWL’s class expressions to define a class based on Concepts’ properties. But the simple approach outlined should be able to cover most of the practical use cases for codelists in agriculture described at the beginning of the article. Applications can then e. g. use the prefLabels of the given concepts in user interfaces for dropdown menus.

7.6. Pulling in labels and relations

There may be cases where ontology developers want to rely on their own class and property definitions but would like to pull in multilingual labels for them from AGROVOC to benefit for search and retrieval or for multilingual user interfaces. Basically, there are two approaches for achieving that:

1. Copying the required labels from AGROVOC by using a SPARQL CONSTRUCT query to build the graph containing the label definitions. Potentially, with that approach label access can be simplified by placing the label literals into rdfs:labels instead of using the full

[16] https://github.com/NatLibFi/Annif
[17] http://uima.apache.org/
[18] https://datenzentrum.bonares.de/research-data.php
[19] http://d2rq.org
[20] https://github.com/antidot/db2triples

SKOS-XL labels used in AGROVOC. That suffices for simple applications that do not require SKOS-XL label features.

2. Using an owl:objectProperty or rdfs:Property to map a class or property definition against an AGROVOC concept. The choice of property is somewhat tricky though: skos:mappingRelation subproperties are unsuited for that, as their rdfs:range and rdfs:domain are defined to be skos:Concepts. They can thus not be used for mapping properties or classes against concepts. owl:sameAs on the other hand is reserved for instances. Therefore, using a custom property might usually be necessary.

7.7. AGROVOC concepts as classes and properties

SKOS allows for skos:Concepts to be redeclared as either classes or properties in an ontology. Statements following these patterns would thus be consistent with the SKOS model:

Redeclare the AGROVOC concept “fertilizers” to be a class in a RDF schema/an ontology:
<http://aims.fao.org/aos/agrovoc/c_2867 > a rdfs:Class.
<http://aims.fao.org/aos/agrovoc/c_2867 > a owl:Class.
Redeclare the AGROVOC concept “soil pH” to be a property in a RDF schema:
<http://aims.fao.org/aos/agrovoc/c_34901 > a rdfs:Property.

If a concept is too broad and unspecific to be used directly in an ontology or schema, subclasses and subproperties can be used to add the required level of detail. Using these design patterns, RDF schemas and ontologies can benefit from the definitions, labels and mappings to other datasets given in AGROVOC.

8. Conclusions

Having successfully piloted the inclusion of specialized concept schemes within AGROVOC, using Vocbench 3, the future looks bright in terms of expanded collaboration with more expert communities. An example might be rural finance: the AGROVOC thesaurus will benefit from new specialized concepts, curated by subject matter experts. At the same time, the community of practice will benefit from AGROVOC infrastructure, the network of international editors, and semantic technology expertise. Partnerships are a cornerstone of this work: AGROVOC relies on the international network of editors and partner institutions, and supporting them remains a key priority. Strengthening alignments to other thesauri is also of high importance, to better support coherence and cohesion in knowledge systems. This will also enhance big data approaches, which facilitate better decision-making and accountability, as well as more effective sharing of knowledge and technologies across the world.

In 2020, the AGROVOC team launched “AGROVOC Massive Open Online Course”. This online course aims to increase awareness of the use of controlled vocabularies to enhance the accessibility and visibility of information and data, with a particular focus on agriculture and related sciences. The objectives are to illustrate the benefits of application of semantic technologies and of using tools like AGROVOC, and to showcase how to effectively and efficiently interact with AGROVOC by equipping participants with the knowledge and skills needed. While AGROVOC is used as an example throughout the course, the content has a wider application for those interested in controlled vocabularies and semantic web technologies.

The AGROVOC team conducts webinars, shares articles on AGROVOC and explains how controlled vocabularies can improve the sharing of information and data. An AGROVOC mailing list is also maintained. This is part of a wider effort by FAO to provide access to current scientific literature and data, and to develop capacity on creating, managing and exchanging open data and research data.

Apart from simple interactive use as a thesaurus, AGROVOC has a number of practical applications and use cases in modern information service infrastructures ranging from text annotation and indexing, applications in research data management to full-fledged integration into ontologies, schemas and data sets.

CRediT authorship contribution statement

Imma Subirats-Coll: Conceptualization, Supervision, Project administration, Funding acquisition, Writing - original draft. **Kristin Kolshus:** Data curation, Conceptualization, Writing - original draft. **Andrea Turbati:** Conceptualization, Methodology, Software, Formal analysis, Writing - review & editing. **Armando Stellato:** Software, Methodology, Formal analysis. **Esther Mietzsch:** Data curation, Writing - review & editing. **Daniel Martini:** Data curation, Writing - original draft. **Marcia Zeng:** Data curation, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work described in this chapter could have not been possible without the collaboration of a number of people. We wish to thank our colleagues Fabrizio Celli, Dongxin Feng, Ilkay Holt, Alejandra Manco Vega, Fernanda Peset, Marina Pines, Giampaolo Rugo, Chelsey Scalese and Karna Wegner.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2020.105965.

References

Stellato, A., Turbati, A., Fiorelli, M., Lorenzetti, T., Costetchi, E., Laaboudi, C., … Keizer, J. (2017). Towards VocBench 3: Pushing Collaborative Development of Thesauri and Ontologies Further Beyond. In P. Mayr, D. Tudhope, K. Golub, C. Wartena, & E. W. De Luca (Ed.), 17th European Networked Knowledge Organization Systems (NKOS) Workshop. Thessaloniki, Greece, September 21st, 2017, (pp. 39-52). Thessaloniki, Greece. Retrieved from http://ceur-ws.org/Vol-1937/paper4.pdf.

Stellato, A., Fiorelli, M., Turbati, A., Lorenzetti, T., van Gemert, W., Dechandon, D., Laaboudi-Spoiden, C., Gerencsér, A., Waniart, A., Costetchi, E., 2020. VocBench 3: a Collaborative Semantic Web Editor for Ontologies, Thesauri and Lexicons. Semantic Web 11, 855–881. IOS Press.

Suominen, O., 2019. Annif: DIY automated subject indexing using multiple algorithms. LIBER Quarter. 29 (1), 1–25. https://doi.org/10.18352/lq.10285.

Musen, M.A, 2015: The Protégé project: A look back and a look forward. AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), June 2015. DOI: 10.1145/2557001.25757003.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., et al., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3 (1). https://doi.org/10.1038/sdata.2016.18.

EPPO Codes. https://www.eppo.int/RESOURCES/eppo_databases/eppo_codes (Accessed: 14 April 2020).

ISOBUS Data Dictionary according to ISO 11783-11. https://www.isobus.net/isobus/ (Accessed: 14 April 2020).

Eurostat: Crop production (apro_cp) Reference Metadata in ESS Standard for Quality Reports Structure. https://ec.europa.eu/eurostat/cache/metadata/EN/apro_cp_esqrs.htm (Accessed: 14 April 2020).

Breed Codes for International Genetic Evaluation of dairy and beef cattle. https://interbull.org/ib/icarbreedcodes (Accessed: 14 April 2020).

Berners-Lee, T., 2006: https://www.w3.org/DesignIssues/LinkedData.html (Accessed: 14 April 2020).

https://www.upov.int/genie/species.xhtml (Accessed: 14 April 2020).

Berners-Lee, T., Fielding, R., Masinter, L., 2005: RFC3986: Uniform Resource Identifier (URI): Generic Syntax. https://tools.ietf.org/html/rfc3986 (Accessed: 14 April 2020).

Schreiber, G., Raimond, Y., 2014.: RDF 1.1 Primer. World Wide Web Consortium Working Group Note. https://www.w3.org/TR/rdf11-primer/ (Accessed: 14 April 2020).

Pazienza, M.T., Scarpato, N., Stellato, A., Turbati, A., 2012. Semantic Turkey: A Browser-Integrated Environment for Knowledge Acquisition and Management. Semantic Web J. 3 (3), 279–292. https://doi.org/10.3233/SW-2011-0033.

Stellato, A, Rajbhandari, S, Turbati, A, Fiorelli, M, Caracciolo, C, Lorenzetti, T, Keizer, J & Pazienza, MT 2015, 'VocBench: a Web Application for Collaborative Development of Multilingual Thesauri', in F Gandon, M Sabou, H Sack, C d'Amato, P Cudré-Mauroux, A Zimmermann (eds.), The Semantic Web. Latest Advances and New Domains (Lecture Notes in Computer Science), Springer, Cham, <https://doi.org/10.1007/978-3-319-18818-8_3>.

Miles, A., Bechhofer, S. (2009): SKOS Simple Knowledge Organization System Reference. World Wide Web Consortium. https://www.w3.org/TR/skos-reference/ (Accessed: 25 March 2020).

W3C OWL Working Group (2012): OWL 2 Web Ontology Language Document Overview (Second Edition). https://www.w3.org/TR/owl2-overview/ (Accessed: 14 April 2020).

Subirats, Imma; Zeng, Marcia L., 2012. Meaningful Bibliographic Metadata (M2B). Rome: Food and Agriculture Organization of United Nations.

Gärtner, P., Svoboda, N., Kühnert, T., Zoarder, M.A., Heinrich, U., 2017. The BonaRes Metadata Schema. BonaRes Series. https://doi.org/10.20387/bonares-5pgg-8yrp.

Hodge, G., 2000. Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files. Council on Library and Information Resources, Washington, DC.

https://public.rma.usda.gov/apps/ReportShare//Extranet/SOB/Listings/CommodityName2014.pdf.

Allemang, D.: 'Sustainability in data and food', Data Intelligence 1(1), p43-57. Available at: https://www.mitpressjournals.org/doi/full/10.1162/dint_a_00005 (Accessed: 18 March 2020).