



A faceted search system for facilitating discovery-driven scientific activities: a use case from functional ecology

Marie-Angélique Laporte, Eric Garnier, Isabelle Mougenot

► To cite this version:

Marie-Angélique Laporte, Eric Garnier, Isabelle Mougenot. A faceted search system for facilitating discovery-driven scientific activities: a use case from functional ecology. first international Workshop on Semantics for Biodiversity, May 2013, Montpellier, France. hal-01577691

HAL Id: hal-01577691

<https://hal.science/hal-01577691>

Submitted on 27 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A faceted search system for facilitating discovery-driven scientific activities: a use case from functional ecology

Marie-Angélique Laporte¹, Eric Garnier², and Isabelle Mougenot¹

¹ UMR 228 ESPACE-DEV, Maison de la Télédétection 34093 Montpellier, France,
`firstname.lastname@ird.fr`

² Centre d'Ecologie Fonctionnelle et Evolutive (UMR 5175), 1919 Route de Mende,
34293 Montpellier Cedex 5, France,
`firstname.lastname@cefe.cnrs.fr`

Abstract. To address biodiversity issues in ecology and assess the consequences of ecosystem changes, large quantities of long-term observational data from multiple data sets need to be integrated and characterized in a unified way. During these last decades, functional trait-based approaches have shown great potential to facilitate the understanding and the prediction of ecosystem changes. To promote data exchange, portability and to drive higher communication between systems, scientific communities are required to acquire data standards. Semantic web (or web of data) provides a realistic solution for these exact requirements. Consequently semantic web allows for creative approaches and offers opportunities to scientists to gain new insight from experimental data. A first step to this goal is to standardize meaningful and precise terms that are interlinked through a dedicated thesaurus that covers the plant functional diversity domain. Therefore this vocabulary can serve as stable reference resources for integration purposes, specifically when published in RDF language and available as linked data on the web. This manuscript presents a web infrastructure, named Thesauform, that fully exploits the key principles of the web of data and its common open data structures in order to guide the plant functional diversity community of experts to build collectively, manage, visualize and query a SKOS thesaurus. A thesaurus dedicated to plant functional traits is used to demonstrate the potential of the approach. Indeed, the thesaurus, built using the Thesauform tool, is used to semantically annotate heterogeneous data sources, such as the TRY database or the Plant Ontology. Then, a faceted search system, based on SKOS collections, enabling thesaurus browsing according to each end-users requirements is expected to greatly enhance the data discovery in the context of biodiversity studies.

Keywords: Tool, Faceted Search, Thesaurus, Semantic annotation, Functional diversity, Web of Data, Plant Trait, Controlled vocabulary, Interoperability, SKOS

1 Introduction

Resolution of key biodiversity issues goes through continued exchanges and co-operation between related domains, such as ecology, taxonomy, genomic, climatology, soil sciences, etc [1], [2], [3]. To address biodiversity issues, it is now widely accepted that a functional approach has strong potential. Indeed, biological traits of organisms have great capabilities to promote a better understanding and to predict global change consequences on the functioning of ecosystems and the services they provide to human societies [4], [5], [6], [7]. A functional trait is defined as: “any morphological, physiological or phenological feature measurable at the individual level, from the cell to the whole-organism level, without reference to the environment or any other level of organization” [8].

Over the last decades, trait-based research has generated huge volumes of data, within multiple contexts of observations and experiments [9]. Considering this, data can be acquired via specific studies and are influenced by peculiar goals. Additionally, these data sets can also be obtained via very different study contexts and are often described in highly specialized terms. Numerous traits can be measured, for instance, on plants [9], [10], [11], [12]. However, data representation and storage do not constitute a major challenge. This is why data generated by functional ecology are only minimally reused or shared within the community, or over communities, mainly due to data heterogeneity. Given these limitations, open web standards and the generation of open web standards for functional ecology would advance the integration of heterogeneous content, with the primary objective of the emergence of new knowledge.

Our primary concern, which focuses on access, sharing and dissemination of information within a community of experts, is oriented towards the semantic web. The web of data [13], [14] provides the concepts, methods and tools, which allow a gradual slide from a web that mostly supports sharing documents to a web that focuses on the sharing of data to ensure their joint and concerted use by software agents. The web of data is primarily based on the key principles of metadata and controlled vocabularies or even ontologies, which should be considered complementary. Thesaurus, which is a type of controlled vocabularies, bypass ambiguity issues in natural language, in order to control and to clarify the access and exchange of information and to facilitate communication. Consequently a thesaurus reflects deliberate choices of communities relatively to the key terms in their expertise field. SKOS (Simple Knowledge Organization System) [15] provides a common format to manage thesaurus adequately. The need for the simultaneous use of multiple vocabularies being increasing in a context of biodiversity studies, SKOS offers not only the mean to build and to publish a thesaurus on the web, but also to anticipate the establishment of cross-references between thesauri. Accordingly, each thesaurus can be considered as a publicly available relevant resource on the web and can be enriched via meaningful navigation between thesauri, when properly described in an adequate format.

In this paper, we present a complete system dedicated to the ecological community allowing it to create, manage, visualize and query a SKOS thesaurus. The final purpose of the thesaurus is to facilitate the integration and the navi-

gation of the information available in multiple data sources. Our previous work focused on how metadata could be exploited during the collaborative building of a thesaurus, through edition and extension mechanisms using the Thesauform tool. The functional plant trait thesaurus (TOP thesaurus, for Trait Of Plant Thesaurus) was built using the Thesauform tool. In this paper, our goal is now to demonstrate the full capabilities of the TOP thesaurus. First, the TOP thesaurus is used to establish mappings between TOP concepts and other data sources, as for instance the TRY database [9] and the Plant Ontology (PO) [16], [17], in a vision of open data sources, in order to both interconnect available information, and semantically annotate data organized into these data sources. Secondly, the TOP thesaurus is exploited through a faceted search engine that reflects end-users interests and preferences, to facilitate the appropriation of the TOP thesaurus by end-users. The facets then act as access points on the interrelated data sources in guiding their navigation. The TOP thesaurus then fully plays its role by supporting the functional plant trait community to manage existing and future datasets and to interconnect them with data from other relevant domains.

This article is organized as follow:

- Section 2 introduces the approach driven with the Thesauform tool to build a functional plant trait thesaurus as a collaborative product. Once the thesaurus has been built, it serves as stable reference resources for integration purposes and it is used to semantically annotate heterogeneous data sources, such as the TRY database or the Plant Ontology.
- Section 3 explains how faceted search enhances the information retrieval. This section gives an overview of the technologies used to query a SKOS thesaurus using end-user preferences.
- Section 4 consists of the implementation of our approach. This section presents the key features of the TOP thesaurus-browsing interface based on faceted search and how this interface is used by functional ecology expert to find relevant information about functional plant traits.
- Finally, section 5 summarizes and discusses the strengths of our approach and refers to future work.

2 Developing a collective thesaurus: the example of the TOP thesaurus

In order to build a collective thesaurus, our recent work focused on the development of a tool, named Thesauform, dedicated to assist domain experts in their task. The Thesauform tool fully relies on semantic web standards, while providing a flexible and user-friendly environment for domain experts [18]. The process of thesaurus co-construction was divided into two phases: (i) an edition phase, during which experts can perform a number of actions in relation to the construction of the vocabulary (addition/deletion of terms and concepts, change of definitions, addition of a commentary, etc.), and (ii) a validation phase, where experts can validate or invalidate the results of the activities completed

during the previous edition phase through a voting procedure. The functional plant trait community has used the Thesauform tool to describe the different functional plant traits in use in the domain.

A part of the TOP thesaurus, based upon the Thesauform, is shown on Table 1. Twenty different experts from the functional plant trait community collaboratively developed the TOP thesaurus. Currently the TOP thesaurus is composed of about 1200 terms regrouped into approximately 1000 concepts. The TOP thesaurus can be used as a bibliographic resource about plant traits information, since it is available as a web resource. For each trait concept, a preferred term, a definition associated to a bibliographic reference and a broader term are provided. In some cases, synonyms (alternative terms), abbreviation, related terms and narrower terms are also specified, as well as a preferred unit³. For instance, the widely used trait “Specific Leaf Area”, also known under the abbreviation SLA, is defined as “the one sided area of a fresh leaf divided by its oven-dry mass” in Cornelissen et al. 2003, and its measurement unit is expressed in meter squared by kilogram of dry mass (m²kg⁻¹[DM]). In the thesaurus, this trait is linked to different other traits. Indeed, it falls under the broader concept of Morphology and it is related to the Leaf Blade Thickness and the Leaf Mass per Area concepts.

The TOP thesaurus serves as a stable reference resource by organizing traits and their information. It extends beyond the users needs by linking information about traits to different available data sources with the purpose of both enriching and facilitating data interpretation, which requires information from different domains. Consequently, TOP thesaurus concepts have been linked to two different data sources, the TRY database and the Plant Ontology (PO). A real advantage of SKOS is to provide properties dedicated to the establishment of cross-references between thesauri. The mapping approaches, on one hand between the TOP thesaurus and the TRY database and on the other hand the TOP thesaurus and PO, rely on the exactMatch and relatedMatch SKOS properties.

The advantage of linking TOP thesaurus concepts to TRY, the biggest functional plant traits database (about 800 traits are measured in TRY on more than 60000 different plant species), is double. First, the mapping TOP/TRY allows TOP thesaurus to unify the access to TRY data, managing the heterogeneity terms used to describe TRY data. Secondly, such a mapping will show the TRY observation number and the geo-referenced observation number for each mapped trait, or the number of different species, on which a given trait has been measured. This information can be useful to account for both the community interest for a given trait or the data available on a trait.

³ The SKOS vocabulary has been expanded to add this information to the TOP thesaurus, considering the importance of the measurement units for trait data interpretation and quality

Table 1. Modified from Laporte et al. 2012. A subset of the traits present in the TOP thesaurus, together with their information attached (definition and associated reference, broader term (if any), narrower term (if any), preferred unit).

| Trait Preferred Label | Définition | Reference of the definition | Synonym Alternative Label | Abbr | Related Term | Broader term (BT) / Narrower term (NT) | Preferred unit |
|-----------------------|--|-----------------------------|---------------------------|------|--|---|--------------------------------------|
| Trait | | | | | | | |
| | Any morphological, physiological or phenological feature measurable at the individual level, from the cell to the whole-organism level | Violle et al., 2007 | | | | NT:Chemical composition, NT:Optical property, NT:Size, NT:Structure, NT:Time related | |
| Specific leaf area | The one sided area of a fresh leaf divided by its oven-dry mass | Cornelissen et al., 2003 | | SLA | Leaf blade thickness Leaf mass per area | BT:Morphology | m ² kg ⁻¹ [DM] |
| Leaf lifespan | The time period during which an individual leaf or part of a leaf is alive and physiologically active | Cornelissen et al., 2003 | Leaf longevity | | | BT:Life cycle | months |
| Specific root length | The ratio of root length to root mass | Cornelissen et al., 2003 | | SRL | | BT:Morphology | m.g ⁻¹ |
| Plant height observed | The shortest distance between the upper boundary of the main photosynthetic tissues on a plant and the ground level | Cornelissen et al., 2003 | Reproductive plant height | | Canopy height | BT:Height, NT:Phanerophytes, NT:Chamaephytes, NT:Hemicryptophytes, NT:Cryptophytes, NT:Geophytes, NT:Therophytes, NT:Helophytes, NT:Hydrophytes | m |

| Trait Preferred Label | Définition | Reference of the definition | Synonym Alternative Label | Abbr | Related Term | Broader term (BT) / Narrower term (NT) | Preferred unit |
|--------------------------|---|-----------------------------|---------------------------|------|--------------------|--|------------------------|
| Leaf mass per area | The oven dry mass of a leaf divided by its one-sided area | Cornelissen et al., 2003 | Leaf specific mass | LMA | Specific leaf area | BT:Morphology | g[DM]m ⁻² |
| Leaf phenology | The timing of foliage of the whole canopy | Cornelissen et al., 2003 | | | | BT:Life cycle | unitless |
| Leaf photosynthetic rate | The photosynthetic rate per unit leaf mass at measurement temperature | Hendry and Grime, 1993 | | JCO2 | | BT:Leaf photosynthetic rate | nmoles /cm2/sec |
| Wood density | The oven-dry mass of a section of a plant's main stem divided by the volume of the same section when still fresh | Cornelissen et al., 2003 | Stem specific density | | | BT:Morphology | kg[DM]dm ⁻³ |
| Bark thickness | Thickness of the part of the stem that is external to the wood including vascular cambium | Cornelissen et al., 2003 | | | | BT:Anatomy | mm |
| Seed dry mass | The air-dried mass of a seed | Cornelissen et al., 2003 | Diaspore mass | SM | | BT:Seed mass | mg |
| Seed shape | The variance of the three dimensions length width and height dividing each dimension by length so length is unity | Thompson et al., 1993 | | | | BT:Morphology | unitless |
| Relative growth rate | The per unit rate if growth of a plant or plant part. | Evans, 1972 | | RGR | | BT:Growth rate, NT:Whole plant relative growth rate, NT:Stem relative growth rate, NT:Root relative growth rate, NT:Leaf relative growth rate, NT:Shoot relative growth rate | g/g/d |

PO, which is a controlled vocabulary describing plant entities, is of great interest for plant traits, since plant traits are measured on plant tissues or organs. The mapping established between TOP thesaurus concepts and PO concepts allow assigning a reference for the plant entities cited in most trait definitions. Moreover, such a mapping approach will be highly beneficial to link data used in ecology or agronomy to data used in genomics. In fact, PO is mainly used by this latter field and provides the opportunity to serve as a first unifying component between the ecological and the genomic world, both of high interest in biodiversity studies.

The TOP thesaurus fulfills its initial role to provide standard vocabulary available to the functional ecology community, and extends beyond the basic needs to ease information retrieval. In this context, a system considering end-user points of view has been developed and offers a faceted search engine.

3 Information retrieval: Faceted search

Information retrieval using free text search is confronted with limitations in terms of accuracy of the result [1], [19]. The use of controlled term and concepts coming from a thesaurus would enhance data queries [3]. Classic semantic search engines based on controlled terms have been widely used to query data in life science fields. Bioportal⁴ [20] is a web portal providing the interrogation of multiple ontologies or controlled vocabularies based on controlled terms. This kind of search mechanism suffers from limitations, since it can be difficult for an unexperienced end-user to find relevant controlled terms. In fact, with classic semantic search engines, most of the time controlled terms are displayed using auto-completed search fields. To overcome this limitation, faceted search engines are an interesting solution as they facilitate the thesaurus appropriation by the end-users. In such search engines, the results are filtered using relevant parameters or categories, each category reflecting the need of users in the thesaurus navigation environment.

On the MUMIA⁵ web site, faceted search (also called faceted navigation or faceted browsing) is defined as: “a technique for accessing a collection of information, allowing users to explore by filtering available information. A faceted classification system allows the assignment of multiple classifications to an object, enabling the classifications to be ordered in multiple ways, rather than in a single, pre-determined, taxonomic order”. Each facet typically corresponds to the common features shared by a set of objects. Faceted searches are commonly used by e-commerce websites to filter the available products based on the parameters most important for the user choice.

Faceted search systems can be applied to SKOS thesaurus. SKOS good practices describe how to represent such a system in a SKOS compliant way [21]. Facets are closely linked to both thesaurus information visualization and thesaurus information restitution, but not to thesaurus structure or to the infor-

⁴ <http://bioportal.bioontology.org/>

⁵ <http://www.mumia-network.eu/index.php/working-groups/wg4>

mation it carries. In thesaurus or in any other controlled vocabulary, concepts can be assembled into semantically meaningful groups, corresponding to facets. Consequently, facets can be defined as skos:collection [21], [22], [23], gathering concepts with common features. For instance, the functional plant trait concept Specific Leaf Area (2) can be grouped with the concepts Leaf Phenology or Leaf Lifespan, because these three concepts share the common feature of being measured on the same plant part, the leaf. But Specific Leaf Area may also be classified with the Xylem Area concept, because these two measurements refer to a size measurement, the area. The categories plant organ and measurement type can then be considered as two access points to query thesaurus. Each user can choose, which access point to use according to own preference.

Faceted search system is so of prior interest to assist users in their information retrieval. Developing such a system based on facets allows taking users interest into account and then to guide dataset consultation. Having data sources semantically annotated with TOP thesaurus concepts can benefit from faceted search engines as well, because thesaurus facets are used as an access point to disseminate information from heterogeneous data sources.

4 Results: approach implementation, user interface

TOP thesaurus trait information will be mainly accessed by experts from the ecology domain. Considering this, we based our work on an user-friendly and easy to use interface, to assist experts in their access and retrieval of pertinent information. In this section we present the key features of our system⁶.

4.1 Semantic search engine

Search is a crucial feature for focused information retrieval. We propose two types of semantic search approaches to access functional plant trait information (cf. 1). First, a classic semantic search engine is available and allows finding traits with controlled trait terms from the TOP thesaurus through an auto-completed field search and a navigation tree. A unique aspect of our work is the implementation of a faceted search engine based on skos collections. This enhanced the semantic search of trait by providing the opportunity to the users to choose his own filters. In Figure 1, end-user selected categories from the available facets (the selected categories are colored in green). The result of such a selection is dynamically updated in the result part.

4.2 External data sources mapping

To address to need of biodiversity studies and to enhance the cooperation and the sharing of heterogeneous data inside the functional plant trait community

⁶ For the features concerning the collaborative thesaurus building using the Thesauform tool, please refer to Laporte et al. 2012.

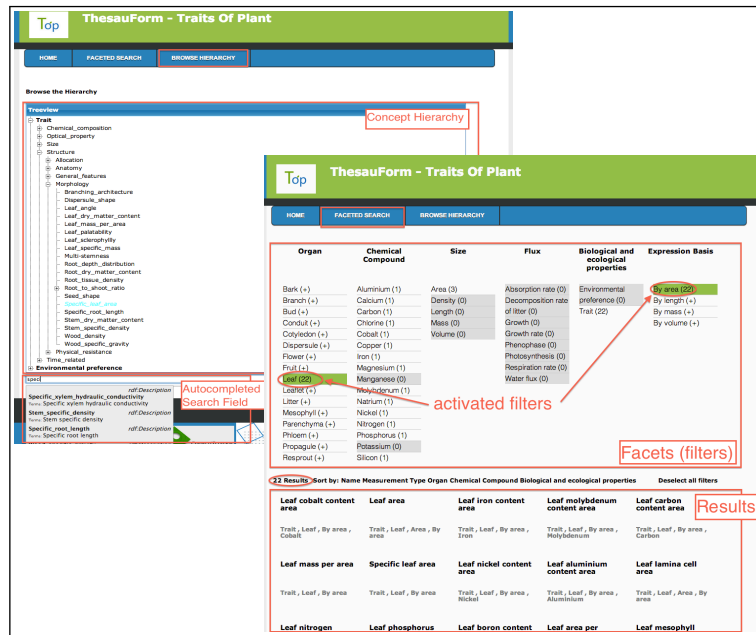


Fig. 1. Semantic search interfaces. On the left, a classic search engine is shown. End-users can use the concept hierarchy tree or an auto-completed field to look for a specific trait. On the right, a faceted search is proposed in order to facilitate and assist the search. The results list is dynamically updated according to the selected filters in the result part.

and over different related domains, specific information for each TOP thesaurus term has been enriched with existing data standard. In Figure 2, the interface displays the results obtained after a query on the TOP thesaurus. The first part of the interface is dedicated to information specific to plant trait, resulting of the collaborative edition of the thesaurus by the community experts. The second part of the interface is about the information from external data sources. For instance, 65157 observations are referenced in the TRY database about Specific Leaf Area.

4.3 Technologies used

We implement a “thin-client/application server” architecture using the J2EE platform, with the system application server being deployed on Apache Tomcat. We used the Jena API to manage the aspects related to the manipulation of the SKOS thesaurus. As we developed a traditional web application, we utilize jquery libraries to support dynamic aspects.

The screenshot shows the 'ThesauForm - Traits Of Plant' interface. At the top, there is a green header with a 'Top' button and the title. Below this is a blue navigation bar with 'HOME', 'FACETED SEARCH', and 'BROWSE HIERARCHY' buttons. The main content area is divided into two sections. The first section, titled 'Specific leaf area', contains the following information: Preferred Name: Specific leaf area; Definition: The one sided area of a fresh leaf divided by its oven-dry mass; Bibliographic Reference: Pérez-Harguindeguy et al, New Handbook; Related Term: Leaf_blade_thickness; Related Term: Leaf_mass_per_area; Preferred Unit: m2kg-1[DM]; Broader Term: Morphology. This section is labeled 'Thesaurus Information'. The second section, titled 'Datasets in TRY database:', contains a table with the following data: TRY trait id: 11; TRY trait name: Specific leaf area (SLA); Observation number in TRY: 65157; Geo-Referenced observation number in TRY: 41479; Species number in TRY: 8612. This section is labeled 'External Information'. Below this, there is a section titled 'Related to PO concepts: leaf', which contains a table with the following data: label: leaf; inSubset: TraitNet; hasExactSynonym: ? (Japanese); onProperty: RO_0002202 someValuesFrom vascular leaf primordium; hasObDef: OBO_SF_PO:3167333; creation_date: 2010-07-12T01:31:44Z; IAO_0000115: A phylome that is not associated with a reproductive structure.; hasOBONamespace: plant_anatomy; id: PO:0025034; created_by: Ramona; onProperty: RO_0002202 someValuesFrom vascular leaf primordium; RO_0002202: PO_0006001. This section is labeled 'Information from TRYDB Plant Ontology (PO)'. At the bottom right, there is a contact information link: 'Contact Marie-Angélique Laporte'.

Fig. 2. Trait restitution information interface. Information from both trait thesaurus and external sources is displayed. By now, information from the TRY Database concerning observation number and type, and plant organs or tissues information from Plant Ontology (PO) have been made available.

5 Conclusion and perspectives

Recent studies highlight the crucial need to dispose thesaurus in the field of biodiversity and more precisely in the field of plant diversity [2], [24]. Plant trait research is complex and requires information from different domains to fully exploit plant trait data. Consequently, we propose a complete system designed to the needs of the plant trait community. Such a system provides a tool to build a SKOS thesaurus, assists a community of experts to manage their datasets, and to interconnect them with data and data standards from related communities using the trait thesaurus. We argue that the end-user preferences have to be of prime importance in data access and retrieval. In this context, a faceted search engine demonstrates its full capabilities. Having data sources semantically annotated with TOP thesaurus concepts can benefit from faceted search engine traits and can be used to access disseminated information from heterogeneous data sources. The approach championed in this paper has been to base our work on the continuity of the Open Linked Data initiative .

The impact of the present work is therefore far reaching. First we propose that, just as the molecular biology community has succeeded in during the past twenty years, the functional ecology community has to widely use controlled

vocabularies, thesaurus and ontology, including the TOP thesaurus, in order to describe and annotate their data in the future years. Second, the available data sets have to be made open source. Third, as illustrated by the use case described in this paper and based on mapping approaches with existing controlled vocabularies or ontologies enhancing data interoperability, the data could reveal their huge capabilities. We highlighted numerous relevant ontologies for such a problematic on the NCBO BioPortal. A next step will be to propose more mapping to external resources (both data and controlled vocabularies/ontologies) with the TOP thesaurus. A significant limitation to this kind of approach in an era of Linked Data is to dispose of controlled vocabularies and ontologies compliant with RDF and all the ensuing Semantic Web standards.

References

1. Jones, M.B., Schildhauer, M.P., Reichman, O., Bowers, S.: The New Bioinformatics: Integrating Ecological Data from the Gene to the Biosphere. *Annual Review of Ecology, Evolution, and Systematics* **37**(1) (December 2006) 519–544
2. Reichman, O.J., Jones, M.B., Schildhauer, M.P.: Challenges and Opportunities of Open Data in Ecology. *Science* **331**(6018) (February 2011) 703–705
3. Michener, W.K., Jones, M.B.: Ecoinformatics: supporting ecology as a data-intensive science. *Trends in ecology & evolution* **27**(2) (February 2012) 85–93
4. Lavorel, S., Garnier, E.: Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology* **16**(5) (October 2002) 545–556
5. Díaz, S., Fargione, J., Chapin, F.S., Tilman, D.: Biodiversity loss threatens human well-being. *PLoS biology* **4**(8) (August 2006) e277
6. Naeem, S., Bunker, D.: TraitNet: furthering biodiversity research through the curation, discovery, and sharing of species trait data. In: *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective*. (2009) 281–289
7. Garnier, E., Navas, M.L.: A trait-based approach to comparative functional plant ecology: concepts, methods and applications for agroecology. A review. *cef-cfr.ca (Umr 5175)* (2011)
8. Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E.: Let the concept of trait be functional! *Oikos* (116) (2007) 882–892
9. Kattge, J., Ogle, K., Bonisch, G., Diaz, S., Lavorel, S., Madin, J., Nadrowski, K., Nollert, S., Sartor, K., Wirth, C.: A generic structure for plant trait databases. *Methods in Ecology & Evolution* (2010)
10. Hendry, G., Grime J.P: *Methods in comparative plant ecology*. Chapman & Hall, London (1993)
11. Cornelissen, J.H.C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D.E., Reich, P.B., Steege, H.T., Morgan, H.D., Heijden, M.G.a.V.D., Pausas, J.G., Poorter, H.: A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany* **51**(4) (2003) 335
12. Knevel, I.C., Bekker, R.M., Bakker, J.P., Kleyer, M.: Life-history traits of the Northwest European flora: The LEDA database. *Journal of Vegetation Science* **14**(4) (2003) 611–614
13. Berners-Lee, T.: Semantic Web. *Conference XML 2000* (2001)

14. T.Berners-Lee: linked data
15. Isaac, A., Summers, E.: SKOS Simple Knowledge Organization System Primer. W3C Technical Report (2008)
16. Walls, R., Cooper, L., Elser, J., Stevenson, D.: The Plant Ontology: A Common Reference Ontology for Plants. wiki.plantontology.org (2010) 2010
17. Cooper, L., Walls, R.L., Elser, J., Gandolfo, M.a., Stevenson, D.W., Smith, B., Preece, J., Athreya, B., Mungall, C.J., Rensing, S., Hiss, M., Lang, D., Reski, R., Berardini, T.Z., Li, D., Huala, E., Schaeffer, M., Menda, N., Arnaud, E., Shrestha, R., Yamazaki, Y., Jaiswal, P.: The plant ontology as a tool for comparative plant anatomy and genomic analyses. *Plant & cell physiology* **54**(2) (February 2013) e1
18. Laporte, M.A., Mougnot, I., Garnier, E.: ThesauForm—Traits: A web based collaborative tool to develop a thesaurus for plant functional diversity research. *Ecological Informatics* (May 2012)
19. Shrestha, R., Arnaud, E., Mauleon, R., Senger, M., Davenport, G.F., Hancock, D., Morrison, N., Bruskewich, R., McLaren, G.: Multifunctional crop trait ontology for breeders' data: field book, annotation, data discovery and semantic enrichment of the literature. *AoB plants* **2010** (January 2010) plq008
20. Noy, N.F., Shah, N.H., Whetzel, P.L., Dai, B., Dorf, M., Griffith, N., Jonquet, C., Rubin, D.L., Storey, M.A., Chute, C.G., Musen, M.a.: BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic acids research* **37**(Web Server issue) (July 2009) W170–3
21. Baker, T., Bechhofer, S., Isaac, A., Miles, A., Schreiber, G., Summers, E.: Key Choices in the Design of Simple Knowledge Organization System (SKOS). eprint [arXiv:1302.1224](https://arxiv.org/abs/1302.1224) (2013)
22. Brugman, H., Malaisé, V., Gazendam, L.: A Web Based General Thesaurus Browser to Support Indexing of Television and Radio Programs. 6–9
23. Suominen, O., Viljanen, K., HyvÄnen, E.: User-Centric Faceted Search for Semantic Portals. In Franconi, E., Kifer, M., May, W., eds.: *The Semantic Web: Research and Applications*. Volume 4519 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg (2007) 356–370
24. Catapano, T., Hobern, D., Lapp, H., Morris, R.A., Morrison, N., Noy, N., Schildhauer, M., Thau, D.: Recommendations for the Use of Knowledge Organisation Systems by GBIF. *Global Biodiversity* (2011)