A Case-Study of Ontology-Driven Semantic Mediation of Flower-Visiting Data from Heterogeneous Data-Stores in Three South African Natural History Collections

Willem Coetzer[1], Deshendran Moodley[2], and Aurona Gerber[3]

CAIR (Centre for Artificial Intelligence Research), University of KwaZulu-Natal
(Durban) / CSIR (Pretoria), South Africa
[1]{w.coetzer@saiab.ac.za} [2]{moodleyd37@ukzn.ac.za}
[3]{agerber@csir.co.za}

Abstract. The domain complexity and structural- and semantic heterogeneity of biodiversity data, as well as idiosyncratic legacy data-creation processes, present significant integration and interoperability challenges. In this paper we describe a case-study of ontology-driven semantic mediation using records of flower-visiting insects from three natural history collections in South Africa. We establish a conceptual domain model for flower-visiting, expressed in an OWL ontology, and use it to semantically enrich the three data-stores. We show how this enrichment allows for the creation of an integrated flower visiting data set. We discuss how this ontology captures both implicit and explicit knowledge, how it can be used to identify and analyze high-level flower-visiting behaviour, and ultimately to construct flower-visiting and pollination networks.

Keywords: biodiversity information, semantic mediation, ontology, plant-insect interactions, pollination

1   Introduction

The challenges of integrating, or making interoperable, distributed, heterogeneous sources of biodiversity- and ecological data have been described [1] [2]. Biodiversity is a complex domain and is no different from other domains in that users encode different definitions of the same concepts [3], which frustrates efforts to integrate data.

We present a case study of three data-stores of flower-visiting insect specimens. All three data-stores consistently contained the names of the plant species, termed *host-plants*, with which both flower-visiting and non-flower-visiting insect specimens were associated. Whereas flower-visiting records were not explicit in most records of two

data-stores, most records of the third data-store contained explicit, easily distinguishable flower-visiting data. To develop a semantic mediation solution, we created the first version of an OWL ontology containing concepts related to flower-visiting and the utilization of flower products, as well as the bearing of pollen by insect vectors. Our work will facilitate the construction of a system to bring about interoperability between distributed and heterogeneous biodiversity data-stores and systems. This will enable biodiversity scientists to more easily extract and analyze the behaviour of flower-visiting insects. Such a system would allow flower-visiting and pollination networks to be automatically assembled and compared.

Outline. In Section 2 we sketch the background against which the need for our study emerged, discuss previous work in biodiversity semantics, and introduce our case-study of interoperability of flower-visiting data. Section 3 begins by describing the domain of flower-visiting and pollination, including our scope, before explaining the process of ontology construction. Expert- and implicit knowledge is highlighted. The usefulness of the concepts in the ontology is discussed in Section 4, by linking data from the data-stores to classes in the ontology. Finally we discuss our approach to a potential solution, including areas where future work is required, and conclude.

2 Background

2.1 Semantics in Biodiversity Informatics

The importance of verifiable specimen-vouchers (i.e. physical preparations such as pinned insects) in museum collections has caused attention to be focused on such specimen information [4]. In recent years *observations* of biodiversity have become important, including observations made by citizen scientists [5]. Both voucher records and observations (collectively termed occurrences) have been subject to the development and adoption of useful standards for publishing and exchanging biodiversity information (the group known as Biodiversity Information Standards (BIS), formerly called the Taxonomic Databases Working Group or TDWG) [6]. One of the BIS standards is the set of terms named the Darwin Core, which contain 'clearly defined semantics that can be understood by people or interpreted by machines, making it possible to determine appropriate uses of the data encoded therein' [7]. The purpose of the Darwin Core terms is to allow biodiversity data to be published and integrated [7].

Biodiversity data are commonly formatted according to the Darwin Core standard and then uploaded to a Global Biodiversity Information Facility (GBIF) participant node (such as the South African Biodiversity Information Facility, SABIF). The data then become discoverable via the GBIF Data Portal, and may be downloaded upon acceptance of conditions. Whereas such database federation has been successful for the sharing of core data attributes (e.g. the Darwin Core categorizes terms as relating to

Occurrence, Event, Location, Identification, Taxon), more specialized data, for example data that record biotic interactions such as parasitism or pollination, are typically omitted because standard terms to describe specific instances of ecological interactions do not yet exist. Currently, shared data therefore fall short of the common phrase 'who did what to whom, where, when, how and why?' because the 'what', 'how' and 'why' are still missing.

The 'Who' and 'To Whom'. The Taxon Concept Schema (TCS) [8] [9], is a standard model to exchange taxonomic information (hence the alternative name 'Taxonomic Concept Transfer Schema'). The TCS is written in XML. More specifically, the TCS allows 'explicit communication of information about Taxon Concepts and their associated names' [8]. A Taxon Concept is a concept or definition of a group, such as a new beetle species, in a taxonomist's mind, which may become published in an article. Several collaborative initiatives aim to define standardized concepts to describe the anatomy and morphology of animals e.g. Hymenoptera [10] or plants [11].

The 'Where' and 'When'. The Darwin-SW Ontology is described as 'an ontology using Darwin Core terms to make it possible to describe biodiversity resources in the Semantic Web' [12]. This is seen as particularly useful for publishing, as Linked Open Data, datasets consisting of Darwin Core terms.

Ecological Semantics. Much work has been done to define concepts used in ecology. Ecological Metadata Language (EML) has a long history of practical application [13] [14], and much work has advanced the use of ontologies [15] [16] to create interoperable systems and to enable the execution of scientific workflows [17] [18].

The need for defining the 'what', 'how' and 'why' of biodiversity information. While the Ecology Ontology and Ecological Networks Ontology [15] contain useful constructs, we found no published, formal definitions of biotic interactions, i.e. concepts that describe specific behaviours representing interactions between individual animals, or between plants and animals. Some preliminary work has been done to extend the Darwin Core standard to broadly include interactions [19] by using terms e.g. `VisitedFlowerOf`, `FlowerVisitedBy`, `NestedIn`, `UsedAsNestBy`. A short list of standard terms was proposed [20] specifically for the interaction, `VisitedFlowerOf`. This list contains the elements: `PollinationEvidence`, `PollenRemoval`, `NectarRemoval`, `OilRemoval` and `FlowerPredation`. Doubt has been expressed as to whether this approach will result in the adequate expression of relationships between specimens or observations.

Semantic mediation in biodiversity informatics. An underlying ontology was used to integrate cereals data from public web databases with data from a local database, allowing molecular characteristics and phenotypic expression to be correlated [37]. While the subject of semantic mediation in biodiversity informatics has been addressed as an architecture component (e.g. [17-18]), few examples of practical applications exist.

2.2 Background to the Case Study

The Quality of Biodiversity Data in South African Museums. South African natural history museums participated in a programme [21] to cleanse and migrate their data to a standard relational database schema and application (Specify Collections Management Software, University of Kansas Biodiversity Institute). Despite having general data of a higher quality, and consistency in schema and syntax, participating researchers of flower-visiting were still unable to easily extract meaningful summaries across data-stores because semantic heterogeneity remained an unresolved challenge. Further work was therefore undertaken with three data-stores that contained data related to collections of flower-visiting insects, namely those of the Albany Museum (AM) in Grahamstown, Iziko Museum (SAM) in Cape Town and Plant Protection Research Institute (SANC) in Pretoria. Table 1 summarizes the data attributes that characterized the data-stores and shows how the word *flower(s)* could be used to distinguish flower-visiting records. The heterogeneity of biodiversity information is evident in Table 1. For example, AM is a specialized flower-visiting data-store because it includes even the colours of visited flowers, and almost all the records are marked with the words 'visit' and 'flower' (also Table 2). On the other hand, SANC contains less-meaningful information for a flower-visiting researcher.

Table 1. Data attributes from the three data-stores. FV = percentage explicit flower-visiting records. Flower-visiting records were distinguished by the *Sampling Method* and *Insect Behaviour* attributes.

	SAM sample data (n=2 094) 3% FV	SANC sample data (n=219) 4% FV	AM sample data (n=21 159) 97% FV
Host Type	host-plant	host-plant	host-plant
Host Taxon	Diascia capensis	Ruschia indecora	Indigofera nigromontana
Sampling Method	**flowers**	swept from **flowering** Acacia albida	hand net
Insect Behaviour	foraging on nectar	[no data]	visiting **flowers**
Flower Colour	[no data]	[no data]	deep pink

3 Ontology Construction in the Domain of Flower-Visiting and Pollination

Various kinds of animals, including arthropods (e.g. insects), birds (e.g. hummingbirds and sunbirds) and mammals (e.g. bats) are well-known *flower-visitors* because they live a life of actively, frequently and consistently seeking out flowers in order to utilize the flowers themselves or their products. The most important flower products are nectar, pollen and oil, which are ingested or collected by the flower-visitors. Insects are important flower-visitors and many insect groups have co-evolved as pollinators of plants.

Pollination is defined with varying granularity. A simple definition reads: 'The transfer of pollen from an anther to a stigma' [22]. Some definitions emphasize that all pollination is ultimately an event (one-step process) because it consists of the act by which pollen is deposited on the pollen-receptive surfaces of a flower (or other reproductive structure such as a cone). In the typical case, pollination (cross-pollination) is a two-step process whereby a vector ('carrier') transfers pollen from the anther of one flower to the stigma of another flower [22]. This is the definition that formed the basis of our domain model, though we did not model the process or event of pollination.

In the study of flower-visiting ecology, pollination may or may not be confirmed in a field setting. Confirmation of pollination requires closely following the flower-visitor and recording its behaviour to see whether it actually transfers pollen onto the stigma. Thus, when ecologists refer to 'pollination' or a 'pollinator', unless otherwise stated, the word is usually used loosely to mean 'inferred pollination' or 'potential pollinator'/'pollen vector' (an organism that carries or transports pollen). Flower-visiting records are the basic currency of pollination ecologists because flower-visiting is easier to observe with high confidence.

Scope. We limited our modelling to angiosperms (flowering plants) that are pollinated by vectors i.e. not by an abiotic medium such as wind or water. We circumscribed as flower-visitors those taxa that belong to the phylum Arthropoda i.e. including the terrestrial groups represented broadly by spiders, millipedes (which mostly inhabit the soil) and insects. Plant galls caused by developing insect larvae, including larvae developing in flower-galls, were excluded from the domain. There was no geographic limitation to our study.

3.1 Concepts used in Domain Modelling: Flower-Visiting and Pollen-Bearing

For the purpose of ontology construction we chose to define the concept of a *flower-visitor* broadly, by interpreting a review of flower-visiting insects [23]. This review clearly included in the concept insects that hid in flowers (e.g. thrips), camouflaged themselves against flowers in order to ambush prey (e.g. mantids) or laid eggs in flowers (e.g. fruit flies). An insect can be a flower-visitor even if it does not ingest or

collect nectar, pollen, oil (with or without terpene fragrance), resin, gum, anthers, ovules, seeds, petals or some other part of the flower or the entire flower.

It is generally accepted that pollen-transfer, both from the anther to a flower-visitor and from the flower-visitor to the stigma is an accidental process.[1] A flower-visitor can become more-or-less covered in pollen, which it may then groom off the surfaces of its body using its tarsi (feet) and mouthparts, and pack into the scopa (hairy patch) on the hind leg, or store on the abdomen or in the crop. The pollen is then taken back to the nest and fed to the young (e.g. social bees) or deposited as nest provision for future young (e.g. solitary bees). Some plants, e.g. orchids and milkweeds, produce a pollinium (plural pollinia), or pollen-mass, borne on a sticky stalk that adheres to the flower-visitor's body. The whole complex including the pollinium and the stalk is called a pollinarium (plural pollinaria).

3.2   Expert- and implicit knowledge

Students of flower-visiting and pollination know implicitly that e.g. an adult beetle or fly or wasp of a certain taxonomic group (e.g. monkey beetles of the tribe Hopliini), or any bee (superfamily Apoidea) has only one reason to be associated with a plant, and that is to visit the plant's flowers, usually to ingest or collect nectar or pollen or other flower products. Many publications list known flower-visiting groups [23].

The importance of implicit knowledge is even more pronounced in the particular case of bees of the genus *Rediviva*, consisting of 26 species that are endemic to South Africa, Lesotho and Swaziland. The females only visit a small number of plant species (about 140 species in 14 genera) whose flowers produce oil to attract these particular bees, or they will visit any number of other plant species whose flowers produce nectar instead of oil [24]. The female bees collect and carry the oil using hairs on their especially-adapted, long front legs, and take the oil back to their nests as provision (i.e. the egg is laid on the oil in the nest and the female that laid the egg then abandons the nest while the larva develops by feeding on the oil). Male *Rediviva* bees only visit flowers that produce nectar, which, like the females that visit 'nectar plants', they ingest to sustain themselves. A 'nectar-plant' could be any flowering plant species, in the area that the bee frequents, that happens to have nectar in its flowers at the time. Among all the specimen records in the SANC data-store that were created during the course of preparing two seminal articles on the famous *Rediviva* oil-collecting bees of southern Africa, the words 'visit', 'flower' or 'oil' do not occur once. The reason for this was probably related to the need for critical information to fit onto a small specimen label. No information was lost within the museum because an expert only needs to know the sex of the adult bee specimen and the plant species name to know whether a *Rediviva* bee was collecting nectar or oil, and that it was visiting flowers[25] [26]).

---

[1] Fig-wasps seem to undertake an intentional pollination ritual [36].

3.3 The Flower-Visiting and Pollen-Bearer Ontology

In this section we describe the semantic analysis and ontology construction process we followed to create the OWL ontology using Protégé [27]. Both bottom-up (i.e. from the data) and top-down ontology construction approaches (i.e. from literature and discussions with experts) were employed. We re-used concepts from the Plant Ontology [11] where possible. In modelling flower-visiting we made extensive use of the `Role` concept as defined in BFO (the Basic Formal Ontology) [28]. Examples of roles include the role of a person as a surgeon or the role of a chemical compound in an experiment. We created –Role concepts for the activities associated with flower visitors , and created an Object Property, `participates_in` (inverse: `participated_in_by`); thus a `FlowerVisitor participates_in` some `FlowerVisitorRole`. The –Role taxonomy is depicted in Figure 1.

Fig 1. The roles (concepts) in the asserted class hierarchy as displayed in Protégé 4.2

3.4 The `FlowerVisitorRole`

Our objective was to make interoperable heterogeneous records of *flower-visitors*, which are generally organisms that utilize flowers. We therefore created the object property, `utilizes` (inverse: `utilized_by`), and defined the necessary condition for the class `FlowerVisitorRole`:

```
utilizes some WholePlant
```
This means that an organism on a severed flower lying on the ground, or in a flower arrangement, cannot be a `FlowerVisitor`.

The necessary and sufficient conditions for the class, `FlowerVisitorRole`, are either:

A: `(utilizes some FlowerMechanicalSupport)`
  `or (utilizes some FlowerSpace)`
  `or (utilizes some FlowerTissue)`
  `or (utilizes some FlowerProduct)`

`or`

B: `(participates_in some PlantVisitorRole)`
  `and (member_of some FlowerVisitingGroup)`

`or`

C: `(bears some Pollen) or (bears some Pollinarium)`

In Section A, `utilizes` some `FlowerMechanicalSupport` could mean alighting on a flower, `utilizes` some FlowerSpace could mean inserting the proboscis into the flower or hiding in the flower. `utilizes` some `FlowerTissue` could mean laying an egg inside the tissue or eating the tissue. `utilizes` some `Flower-Product` could mean ingesting or collecting nectar or pollen. This class will therefore include individuals that are incidental flower-visitors (e.g. spiders) as well as highly specialized pollen-collectors (e.g. bees).

Section B in the above class definition states that a condition for an organism that `participates_in` the `FlowerVisitorRole` is that it `utilizes some WholePlant and is a (member_of some FlowerVisitingGroup)`.

We created the object property, `bears` (inverse: `borne_by`), meaning to 'have on (the outside of the body)', as in 'the bee's abdomen bears pollen'. This object property was used, in Section C above, to assert that a condition for an organism that `participates_in` the `FlowerVisitorRole` is that it bears `Pollen` or bears at least one `Pollinarium`.

3.5 The `FlowerUtilizerRole` and descendent classes, including implicit knowledge of *Rediviva* bees

It was asserted that a condition for the `FlowerUtilizerRole` is ((`utilizes some FlowerMechanicalSupport) or (utilizes some FlowerSpace)or (utilizes some FlowerTissue) or (utilizes some FlowerProduct)`). This means that `FlowerUtilizerRole` is equivalent to `FlowerVisitorRole`.

We specialized the object property, `utilizes`, into the object properties, `ingests` (inverse: `ingested_by`) and `collects` (inverse: `collected_by`).

We defined a `FlowerProduct` to be the class subsuming the class (`FlowerSecretion` or `Pollen` or `Pollinarium`). The class `FlowerSecretion` subsumed the class (`FlowerGum` or `FlowerNectar` or `FlowerOil` or `FlowerResin`).

The `FlowerUtilizerRole` was specialized into `FlowerProductUtilizerRole` and `FlowerPollenBearerRole`. More specifically, if an individual `utilizes` (`ingests` or `collects`) some `FlowerProduct`, that is sufficient to mean that it `participates_in` the `FlowerProductUtilizerRole`.

An individual that (`bears some Pollen`) or (`bears some Pollinarium`) sufficiently meets the condition for the `FlowerPollenBearerRole`. If an organism actively ingests or collects pollen, some pollen will invariably remain on its body after grooming and packing into the scopa. A necessary condition of the `FlowerPollenIngestorRole` and the `FlowerPollenCollectorRole` is therefore: `bears some Pollen`. Figure 2 depicts two parts of the inferred class hierarchy: `FlowerProductUtilizer` and sub-classes, as well as detail of the `FlowerPollenCollector` class hierarchy. The classes in Figure 2 are sub-classes of `Organism`. These classes `participate_in` the `-Role` classes depicted in the taxonomy in Figure 1.
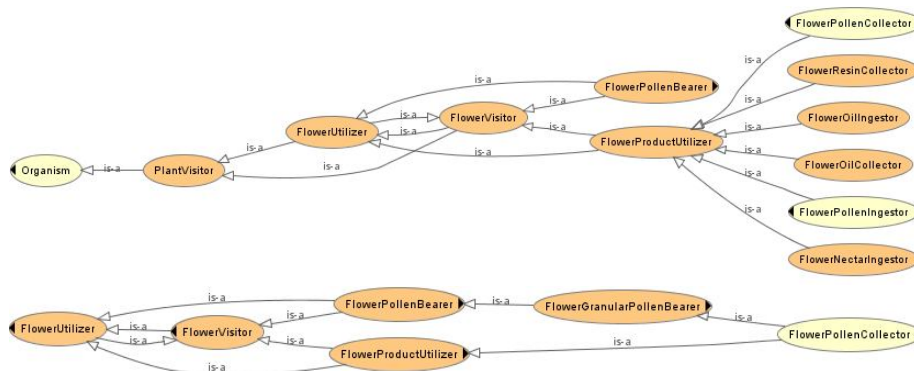

Fig. 2. It is asserted that a `FlowerPollenBearer` need not be a `FlowerProductUtilizer`, but an organism may be both a `FlowerPollenBearer` and a `FlowerProductUtilizer` because these classes are not disjoint. This successfully models active pollen-collecting and pollen-ingesting, which necessarily result in passively bearing pollen.

The conditions that are sufficient for membership in the `FlowerOilCollector` class are as follows: ((`participates_in some FlowerOilCollectorRole`)) or ((`participates_in some OilPlantVisitorRole`) and

```
(member_of some FlowerVisitingGroup) and (has_sex only
Female) and (part_of some RedivivaGenus)).
```

This means that a `FlowerOilCollector` can either be observed directly (`col-lects some FlowerOil`) or its presence can be inferred (e.g. in the SAM data-store) from the facts that an 'oil plant' (with flowers that secrete oil, not nectar) was visited, the insect was a female and it was a species in the genus *Rediviva*.

3.6 The `IllegitimateFlowerVisitorRole` and sub-classes

With reference to Figure 1, the concept of 'illegitimately' visiting flowers (i.e. by definitely avoiding coming into contact with the anthers, and therefore never becoming a `FlowerPollenBearer`) is frequently encountered in the flower-visiting literature, and we therefore included this in our ontology. Robbers, which damage the petals (e.g. by biting a hole in the petal to access the nectar), are distinguished from thieves, which inflict no petal damage. A secondary robber obtains nectar through the hole made by a primary robber [29].

4 Linking the Ontology to Existing and Future Data

The class, `FlowerUtilizer` (Section A of the definition of the `FlowerVisi-torRole`) therefore represents records resulting from the observations of a generalist scientist who may record an organism generally utilizing a flower by e.g. sitting on, or flying around and feeding from (visiting), a flower. In the AM data-store a small number of records were classified as members of the class `FlowerUtilizer` (Table 2).

Table 2. Examples of the class FlowerProductUtilizer in the AM data-store

# records	Behaviour	Class
137	Visiting extrafloral nectaries	PlantVisitor
95	On foliage	PlantVisitor
8	On stem of plant	PlantVisitor
20135	Visiting flowers	FlowerProductUtilizer
380	In flowers	FlowerUtilizer
22	On flowers	FlowerUtilizer
16	Sheltering in flower	FlowerUtilizer
8	In copula on flowers	FlowerUtilizer

The vast majority of records, however, were instances of the class, `Flower-ProductUtilizer`. An expert in the study of flower-visitors would record a flow-er-visitor to be an instance of the class `FlowerProductUtilizer` (i.e. specifical-ly ingesting or collecting nectar or pollen). Importantly, this observation can be made by an expert observing an insect that has not even touched a flower. The expert is able

to classify the organism into a specific taxonomic group, and to remember how previous individuals in this specific group have behaved (i.e. they *visited* flowers, which is a shorter way of recording that they ingested or collected nectar or pollen), and to know that newly observed individuals of the same group are unlikely to behave differently. The predominance of records of the `FlowerProductUtilizer` class therefore reflects the predominance of bees and pollen wasps in this data-store, which is due, in turn, to the development of the careers of the specialists who built the specimen collection. It is therefore not surprising that the biodiversity information in the AM data-store is richer than the information in the other data-stores.

4.2   Data in the SAM and SANC data-stores

Ninety-seven per cent of the records in the SAM data-store, and 96% of the records in the SANC data-store, were instances of the class `FlowerVisitor`, a term that is less meaningful than `FlowerUtilizer` or `FlowerProductUtilizer`. A small number of records in the SAM data-store were instances of sub-classes of the class `FlowerProductUtilizer`. Some of these are shown in Table 3.

Table 3. Examples of the class FlowerProductUtilizer in the SAM data-store

# records	Behaviour	Class
1	Collecting pollen on yellow flowers.	FlowerPollenCollector
1	Patrolling Corymbium. With pollenaria.	FlowerPollinariumBearer
1	Feeding on Brunia laevis pollen	FlowerPollenIngestor
1	Foraging on nectar of Euphorbia flowers.	FlowerNectarIngestor
1	Taking resin from Dalechampia capensis.	FlowerResinCollector

Section C of the definition of the `FlowerVisitorRole` (i.e. a `FlowerPollen-Bearer`) is of particular, current interest. If an organism is seen to bear pollen or a pollinarium, DNA barcoding can be used to identify [30] the plant species that produced the pollen. This is a very important step in the study of flower-visiting because it means that it will no longer be necessary to observe a `FlowerPollenBearer`, either in any physical association with a plant or flower, or actually ingesting or collecting pollen, to know:

1)   That it must be a `FlowerUtilizer` (but not necessarily a `Flower-ProductUtilizer`) and therefore a `FlowerVisitor`;
2)   The list of plant species which it has recently visited, utilized and borne pollen from.

5 Discussion and Conclusion

We have shown how implicit domain knowledge about flower visitors can be represented in an ontology for use in semantic enrichment of, and semantic mediation between, heterogeneous data sources.

Researchers of flower-visiting need to summarize data into lists of insect species and the plant species whose flowers those insects visit, and which they probably pollinate. These lists usually form the basis of further work involving the modelling of flower-visiting networks (which are useful in community ecology), and, more specifically, pollination networks (e.g. [31]). In an applied study the ultimate objective may be to compare the characteristics [32] of pollination networks across space or through time e.g. to estimate the effect, on pollination, of habitat transformation [33] or global change.

Clearly, systems used to capture and manage specimen data are not designed to capture the background knowledge required to access the rich, and often implicit, information associated with these records. This knowledge is usually held by the curator or scientists who generated the records. This becomes more pronounced for biodiversity researchers accessing a network of locally controlled and heterogeneous biodiversity databases. A significant barrier to data integration and analysis will therefore be removed if knowledge can be explicitly represented within the system. For example, illegitimate flower-visitor species must be excluded from the process of assembling a pollination network.

In our current ontology we assumed that there are no exceptions of a `Known-FlowerVisitingGroup`. This is an area where future work is needed because the semantic representation of exceptions, or defeasibility with current OWL ontologies, is problematic. One of these exceptions is a particular Afrotropical bee species which is an obligate raider of other bees' nests and therefore has no need to, and never does, visit flowers. Yet bees are the most important group of flower-visiting insects. Such exceptions will need to be carefully modelled to prevent the possibility of drawing incorrect inferences.

While the ontology described above can certainly facilitate the creation of a semantically rich flower-visiting data set, it still falls short of capturing uncertain and vague biotic interactions associated with flower-visiting occurrences. Probabilistic graphs such as Bayesian Networks are better able to deal with uncertain causal relations, especially when there is uncertainty and vagueness [34]. The combination of ontologies and Bayesian networks has recently been explored in the earth observation domain within the Sensor Web Agent Platform (SWAP) [35]. In SWAP sensor observations from heterogeneous sensor data-stores are semantically enriched with OWL ontologies and used to populate Bayesian networks to determine the probability of the occurrence of abstract physical earth observation phenomena.

The next step in our semantic mediation system will be to adapt the SWAP [35] approach and construct a Bayesian network that describes the causal relations between plant-visiting events, flower-visiting events, pollen transfer events and pollination events. These events will be defined using concepts from the flower-visiting ontology. In this way semantically enriched observations from the three data-stores can be used as proxies to determine the probabilities of the occurrence of flower-visiting and pollination events.

Acknowledgement

References

1. Johnson, N.F.: Biodiversity Informatics. Annual Review of Entomology. 52, 421–38 (2007).
2. Jones, M.B., Schildhauer, M.P., Reichman, O.J., Bowers, S.: The New Bioinformatics: Integrating Ecological Data From the Gene to the Biosphere. Annual Review of Ecology Evolution and Systematics. 37, 519–544 (2006).
3. Deans, A.R., Yoder, M.J., Balhoff, J.P.: Time to Change How We Describe Biodiversity. Trends in Ecology & Evolution. 27, 78–84 (2011).
4. Bisby, F.A.: The Quiet Revolution: Biodiversity Informatics and the Internet. Science. 289, 2309–2312 (2000).
5. Silvertown, J.: A New Dawn For Citizen Science. Trends in Ecology & Evolution. 24, 467–471 (2009).
6. Biodiversity Information Standards, http://www.tdwg.org/.
7. Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., Vieglais, D.: Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE. 7, e29715 (2012).
8. Kennedy, J., Hyam, R., Kukla, R., Paterson, T.: A Standard Data Model Representation for Taxonomic Information. Omics, a Journal of Integrative Biology. 10, 220–230 (2006).
9. Hyam, R., Kennedy, J.: Taxon Concept Schema – User Guide. Unpublished Report. 28 pp. (2005).
10. Yoder, M.J., Mikó, I., Seltmann, K.C., Bertone, M.A., Deans, A.R.: A Gross Anatomy Ontology For Hymenoptera. PloS one. 5, e15991 (2010).
11. The Plant Ontology Consortium: The Plant Ontology^{TM} Consortium and Plant Ontologies. Comparative and Functional Genomics. 3, 137–142 (2002).
12. Webb, C., Baskauf, S.: Darwin-SW: Darwin Core Data for the Semantic Web.
13. Michener, W.K., Brunt, J.W., Helly, J.J., Kirchner, T.B., Stafford, S.G.: Nongeospatial Metadata for the Ecological Sciences. Ecological Applications. 7, 330–342 (1997).

14.   Johnson, J.C., Christian, R.R., Brunt, J.W., Hickman, C.R., Waide, R.B.: Evolution of Collaboration within the US Long Term Ecological Research Network. BioScience. 60, 931–940 (2010).

15.   Williams, J.R., Martinez, N.D., Golbeck, J.: Ontologies for Ecoinformatics. Web Semantics: Science, Services and Agents on the World Wide Web. 4, 237–276 (2006).

16.   Madin, J., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D., Villa, F.: An Ontology for Describing and Synthesizing Ecological Observation Data. Ecological Informatics. 2, 279–296 (2007).

17.   Michener, W.K., Beach, J.H., Jones, M.B., Ludäscher, B., Pennington, D.D., Pereira, R.S., Rajasekar, A., Schildhauer, M.: A Knowledge Environment for the Biodiversity and Ecological Sciences. Journal of Intelligent Information Systems. 29, 111–126 (2007).

18.   Michener, W.K., Jones, M.B.: Ecoinformatics: Supporting Ecology as a Data-Intensive Science. Trends in Ecology & Evolution. 27, 85–93 (2012).

19.   De Giovanni, R., Cartolano, E., Giannini, T., Saraiva, A., Pizzigatti, P.: Darwin Core Interaction Extension Concept List, http://wiki.tdwg.org/twiki/bin/view/DarwinCore/InteractionExtension.

20.   De Giovanni, R., Cartolano, E., Giannini, T., Saraiva, A., Pizzigatti, P.: Darwin Core Interaction Extension: Pollination Extension Concept List, http://wiki.tdwg.org/twiki/bin/view/DarwinCore/PollinationExtension.

21.   Coetzer, W., Gon, O., Hamer, M., Parker-Allie, F.: A New Era for Specimen Databases and Biodiversity Information Management in South Africa. Biodiversity Informatics. 8, 1–11 (2012).

22.   Raven, P.H., Evert, R.F., Eichhorn, S.E.: Biology of Plants. Worth Publishers, Inc., New York (1986).

23.   Kevan, P.G., Baker, H.G.: Insects as Flower Visitors and Pollinators. Annual Review of Entomology. 28, 407–453 (1983).

24.   Pauw, A.: Floral Syndromes Accurately Predict Pollination by a Specialized Oil-Collecting Bee (Rediviva peringueyi, Melittidae) in a Guild of South African Orchids (Coryciinae). American Journal of Botany. 93, 917–926 ST – Floral syndromes accurately predict (2006).

25.   B Whitehead, V., E Steiner, K.: Oil-collecting Bees of the Winter Rainfall Area of South Africa. Annals of The South African Museum. 108, 143–277 (2000).

26.   Whitehead, V.B., Steiner, K.E., Eardley, C.D.: Oil Collecting Bees Mostly of the Summer Rainfall area of Southern Africa (Hymenoptera: Melittidae: Rediviva). Journal of the Kansas Entomological Society. 81, 122–141 (2008).

27.   Horridge, M.: A Practical Guide To Building OWL Ontologies Using Protege 4 and CO-ODE Tools Edition 1.3, (2011).

28.   Arp, R., Smith, B.: Function, Role, and Disposition in Basic Formal Ontology. Nature. 2, 1–4 (2008).

29.   Murphy, C.M., Breed, M.D.: Nectar and Resin Robbing in Stingless Bees. American Entomologist. Spring, 36–44 (2008).

30.   Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R.: Biological identifications through DNA Barcodes. Proceedings of the Royal Society B: Biological Sciences. 270, 313–321 (2003).

31. Dupont, Y.L., Padrón, B., Olesen, J.M., Petanidou, T.: Spatio-Temporal Variation in the Structure of Pollination Networks. Oikos. 118, 1261–1269 (2009).

32. Kaiser-Bunbury, C.N., Muff, S., Memmott, J., Müller, C.B., Caflisch, A.: The Robustness of Pollination Networks to the Loss of Species and Interactions: A Quantitative Approach Incorporating Pollinator Behaviour. Ecology Letters. 13, 442–452 (2010).

33. Valdovinos, F.S., Ramos-Jiliberto, R., Flores, J.D., Espinoza, C., López, G.: Structure and Dynamics of Pollination Networks: The Role of Alien Plants. Oikos. 118, 1190–1200 (2009).

34. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs, NJ (2003).

35. Moodley, D., Simonis, I., Tapamo, J.: An Architecture for Managing Knowledge and System Dynamism in the Worldwide Sensor Web. International Journal of Semantic Web and Information Systems: Special issue on Semantics-enhanced Sensor Networks. Internet of Things and Smart Devices. 8, 64–88 (2012).

36. Wiebes, J.T.: Co-evolution of Figs and Their Insect Pollinators. Annual Review of Ecology and Systematics. 10, 1–12 (1979).

37. Sala, A., Bergamaschi, S.: A Mediator Based Approach to Ontology Generation and Querying of Molecular and Phenotypic Cereals Data. International Journal of Metadata, Semantics and Ontologies. 4(1/2), 85-92 (2009).