

GoodRelations: An Ontology for Describing Products and Services Offers on the Web

Martin Hepp

E-Business and Web Science Research Group, Bundeswehr University Munich, Germany
mhepp@computer.org

Abstract. A promising application domain for Semantic Web technology is the annotation of products and services offerings on the Web so that consumers and enterprises can search for suitable suppliers using products and services ontologies. While there has been substantial progress in developing ontologies for *types* of products and services, namely eClassOWL, this alone does not provide the representational means required for e-commerce on the Semantic Web. Particularly missing is an ontology that allows describing the relationships between (1) Web resources, (2) offerings made by means of those Web resources, (3) legal entities, (4) prices, (5) terms and conditions, and the aforementioned ontologies for products and services (6). For example, we must be able to say that a particular Web site describes an offer to sell cell phones of a certain make and model at a certain price, that a piano house offers maintenance for pianos that weigh less than 150 kg, or that a car rental company leases out cars of a certain make and model from a set of branches across the country. In this paper, we analyze the complexity of product description on the Semantic Web and define the GoodRelations ontology that covers the representational needs of typical business scenarios for commodity products and services.

Keywords: Semantic Web, E-Commerce, E-Procurement, eClassOWL, UNSPSC, eCl@ss.

1 Overview

There is a vast body of work in the Semantic Web and ontology research communities that stresses the potential of using ontologies for improving the search for products on the Web or for managing product data in B2B e-commerce scenarios, e.g. [1-3]. Surprisingly, nobody has yet tried to exploit this potential and develop the ontologies and infrastructure for real Semantic Web-based e-commerce on a Web scale – compliant with current Semantic Web standards and deployable on available Semantic Web infrastructure. In this paper, we aim at filling that gap.

The semantics of representing products and services themselves is now pretty well understood, and there exist prototypes from controlled B2B settings [4-6] and general products and services ontologies in the official ontology languages for the Semantic Web, namely eClassOWL [7] and a similar ontology unspscOWL, which is awaiting copyright clearance. However, all of these components alone do not yet provide the means required for e-commerce on the Semantic Web. This is because annotating

offerings on the Web requires much more complex statements than “Resource A is an instance of the product class TV set”. Foremost, we need to express the business relation between such a resource and the legal entity making that offering: Are they offering to *sell* this instance, or do they offer it for rent? For which audience is this offer intended and valid, both in terms of eligible regions (Austria and Germany only, all of Europe, worldwide, ...) and who are eligible types of buyers (end users, wholesalers, ...)? Also, we need to be able specify a wealth of terms and conditions, and we must clearly differentiate between information resources representing the description of an offering, and the offering itself.

In short, we currently lack a non-toy ontology that allows describing the *relationship* between (1) Web resources, (2) offerings made by means of those Web resources, (3) legal entities, (4) prices, (5) terms and conditions, and the aforementioned ontologies for products and services (6).

In this paper, we analyze the complexity of product description in the Semantic Web and propose the *GoodRelations* ontology that covers the representational needs of typical e-commerce scenarios in the commodity segment. It is part of the myOntology and ebSemantics initiatives, in which the infrastructure for Semantic Web-based e-commerce for Austria is being developed.

Due to space constraints, this submission cannot present all technical details of our work. A comprehensive Technical Report and the GoodRelations ontology are available at <http://www.heppnetz.de/projects/goodrelations/>.

2 Representational Requirements

In this section, we develop the requirements on an ontology for commodities and commodity services offered on the Web.

2.1 Motivating Scenarios

The following are typical examples of offerings made on the Web.

Scenario 1: A Web resource represents an entity that, *in general*, offers items of a particular kind of good for sale, either to wholesalers or to end users, or both; they might offer concrete, identifiable instances or it may be that it is only said that such instances exist.

Scenario 2: A Web resource describes *the make and model of a commodity* and its specifications. Such Web resources are usually within the domain name space of the respective manufacturer. There may exist actual instances of this make and model, which by default inherit their properties from the specifications of the make and model, but also have additional properties of their own (e.g. the date of production, the serial number).

Scenario 3: A Web resource represents (1) the description of a particular range of products, determined either by product classes or makes and models, and property ranges, and (2) a concrete offer *to rent out* unidentified instances thereof. Such Web resources are usually within the domain name space of rental agencies or Web pages of local dealers.

Scenario 4: A particular Web resource represents (1) the description of a particular range of products, determined either by product classes, makes and models, or property ranges, and (2) a concrete offer *to provide a certain type of service* for this range of products (e.g. maintenance, repair, or disposal). Such Web resources are usually within the domain name space of local dealers.

2.2 Competency Questions

Next, we specify the scope and purpose of the GoodRelations ontology using *competency questions*, which is a standard technique in ontology engineering methodologies [8]. The questions were discussed with stakeholders, namely the developers of recommender systems, operators of Web shops, and other domain experts.

CQ1: Which retrievable Web Resources describe an offer

- {to sell | to provide the service of | to repair | to maintain | to lease out | to dispose}
- {a concrete individual | some unknown individuals} of
- a {given good | given service | spare part for a given good | consumables and supplies for a given good} described by a {type of good | specific make and model}
- that meet certain requirements on {properties | intervals for properties}
- for which the offering party accepts a given method of payment and
- provides a certain method of delivery
- to {consumers | retailers}
- in a given {country | region}?

CQ2: For which time frame is the offer valid?

CQ3: Which types of customers are eligible?

CQ4: Which are the eligible customer regions?

CQ5: Which shipping / delivery methods are available?

CQ6: Which methods of payment are accepted?

CQ7: For any such offer, what is the price and currency for a given quantity, delivery region, and type of customer, per unit of measurement? Does the price include VAT and sales taxes?

CQ8: What is the shipping charge and currency for a given delivery method to a given region? Does it include VAT?

CQ9: What is the payment charge and currency for a given payment method? Does it include VAT?

CQ10: What is the mail address and which are the contact details of the offering business entity?

CQ11: Which are the locations from which the product or service is being provided, what are the contact details and opening hours of each location?

CQ12: What is the scope and duration of the warranty promise or warranty promises included in the offer?

CQ13: Which offerings on the Web refer to {spare parts | consumables or supplies} for a given {type of good | make and model}

3 Domain Capture

In the following, we give an overview of the relevant conceptual entities and types of relationships. A visualization of our domain capture in the form of an entity-relationship diagram is shown in Figure 1. Note that certain limitations of OWL, namely the lacking support for relations with a higher arity than two, require modeling workarounds that introduce new conceptual elements, which are only introduced in the later ontology coding stage. We use **bold** characters when introducing and defining a conceptual element, and underlining to refer to the particular definition of a word or group of words defined elsewhere.

The most important conceptual elements of the domain are as follows:

Web Resource: A retrievable Web resource that contains information related to a business entity, an offering, a product model, or similar. Very often, a Web Resource combines information about multiple conceptual entities, e.g. about the offering party and the types of products. In the GoodRelations ontology, we assume Web Resources to be instances of `rdfs:Resource` and use the `rdfs:seeAlso` property for linking the actual conceptual entities (business entities, offerings, product instances, etc.) to Web Resources that contain a human-readable description. This is why there is no class “Web Resource” in the GoodRelations ontology.

Business Entity: A legal agent making a particular offering. This can be a legal body or a person. A Business Entity has at least a primary mailing address and contact details. For this, typical address standards (vCard) and location data can be attached. (Example: Siemens Austria AG, Volkswagen Ltd., Peter Miller's Cell phone Shop).

Offering: The public, not necessarily binding, not necessarily exclusive, announcement by a Business Entity to provide a certain Business Function for a certain Product or Service Instance to a particular target audience. An Offering is specified by the type of product or service or bundle it refers to, what Business Function is being offered (sales, rental, ...), and a set of commercial properties. It can either refer to a clearly specified instance (Product or Service Instance) or to a set of anonymous instances of a given type (existentially quantified Product or Service Instances, see below). An offering may be constrained in terms of the eligible type of business partner, countries, quantities, and other commercial properties. The definition of the commercial properties, the type of product offered, and the business function are explained in the following sections in more detail. (Example: Peter Miller offers to repair TV sets made by Siemens, Volkswagen Innsbruck sells a particular instance of a Volkswagen Golf at \$10,000.).

Business Function: The type of activity or access offered by the Business Entity on the Product or Services through the Offering. The idea of standardizing business functions was first put to practice by the UNSPSC Business Functions Identifiers [9]. The

Product or Service: Instance, Model, and Class: In the products and services domain, we find multiple types of conceptual entities when it comes to describing what is being offered: First, actual products, like for example my cell phone or a concrete TV set. Second, certain product makes and models, e.g. the cell phone make and model Sony 1234 or the car model Ford T. There usually exist actual products that are of the respective make and model, but they all have an identity of their own. In particular, they differ in several properties. Third, classes of actual products that are similar in function or nature, like for example the class “cell phone” which subsumes all actual cell phones. These elements are discussed in more detail in section 3.1.

Property of a Product or Service: Products usually have certain characteristic features, often physical or chemical attributes or such related to the context of usage. A Product or Service Property is the type of a characteristic feature of an actual product or service instance. The value may be a quantitative or a qualitative value. In the former case, a Product or Service Property is a *ternary* relation between a Product or Services Instance, a Unit of Measurement, and this value. In the latter case, it is a binary relation between a Product or Services Instance and this value. A quantitative value is a numerical interval that represents the range of a certain quantitative feature in terms of the lower and upper bounds for one particular. It is to be interpreted in combination with the respective Unit Of Measurement. Most quantitative values are intervals even if they are in practice often treated as a single point. A qualitative value is an entity that represents the state of a certain qualitative feature and can be either a literal value or an enumerative value. Products or Services Classes may be defined by constraints on Product or Services Properties. Product or Services Models may provide default values for Product or Services Properties.

Unit of Measurement: The point of reference and the scale for quantitative values. Those can refer to the technical/physical (e.g. weight) or commercial perspectives (e.g. sales units). Work on standardizing Units of Measurement has been ongoing in various standardization bodies, but there is still no single, widely agreed standard in place. A common approach is the UN/CEFACT Recommendation No. 20 [10], which we adopt for GoodRelations.

Commercial Property of the Offering: An offering is usually more specific than just stating the general availability of a given type of product or service. In many cases, it is desirable to specify additional commercial aspects, first of all the price. Commercial properties are discussed in more detail in section 3.2.

Relationships between Multiple Product Models or Product Instances: In the commodities sector, it is often valuable to express several relationships between Product Instances and/or Product Models. GoodRelations defines “isConsumableFor”, “isSimilarTo”, and “isAccessoryOrSparePartFor” as the most common ones. For the full details, please see the Technical Report that complements this paper.

3.1 Products or Services: Instances, Models, and Classes

As explained, there are three conceptual elements related to the type of product or service being offered, which we specify in the following. In principle, we find the

same pattern in commodity services¹ as well. For example, there is the class “haircutting” (a subclass of “Happening”) which subsumes all instances of someone actually cutting someone else’s hair. In some cases, there may also be makes and models for services, but this is not very common in our opinion.

Product or Service Instance: A single identifiable object or action that creates some increase in utility (in the economic sense) for the individual possessing or using this very object (Product) or for the individual in whose favor this very action is being taken (Service). Products or Services are types of goods in the economic sense. For an overview of goods and commodities in economics, see [11]. (Examples: MyThinkpad T60, the pint of beer standing in front of me, my Volkswagen Golf, the haircut that I received or will be receiving at a given date and time.) Note that in many cases, product or service instances are not explicitly exposed on the Web but only existentially quantified. For a detailed discussion and practical solutions, see section 4.3.

Product or Service Model: An intangible entity that specifies some characteristics of a group of similar, usually mass-produced Products. In case of mass-produced Products, there exists a relation hasMakeAndModel between the Products and Services Instance and the Product or Service Model. A Product or Service Model may carry information on characteristics or features of actual products of this make and model. For example, a certain camera make and model may specify the default weight of an actual camera. It is very likely that an actual camera of that make and model will show that weight; however, it is not guaranteed. If we cut off a piece of that camera, the camera will have a different weight but still be a camera of the given make and model. (Examples: Ford T, ThinkPad T60, Volkswagen Golf)

Note that from the ontological perspective, Product or Service Models are not subclasses of Products or Services, but intangible objects in their own right. The main reason is that the model has an identity of its own. For a formal justification of that distinction, please refer to the OntoClean methodology [12, 13].

Product or Service Class: A set of Product or Services Instances that provide the same functionality, can be used for achieving the same goal, or are similar in their physical characteristics. (Examples: Laptop computer, car, writing paper, hot beverage, or haircutting as a service.)

3.2 Commercial Properties of the Offering

In many cases, it is desirable to specify additional commercial aspects than just the fact that a certain good is being offered. Related work to this issue has been done in the context of XML formats for catalog data exchange. Kelkar et al. [14], for example, list prices, quantities, eligible regions, and eligible types of business partners as relevant details. In the following, we describe the relevant elements for the GoodRelations ontology.

Price Specification: The price asked for a given Offering by the respective Business Entity. An Offering may be linked to multiple Price Specifications that specify

¹ We are talking about *real* services here, not Web services.

alternative prices for non-overlapping sets of conditions (e.g. quantities or sales regions). For GoodRelations, we take the price model from Kelkar, Leukel, and Schmitz (2002, [14]) as a starting point and reuse a subset there from. A Price Specification is characterized by (1) the lower and upper limits and the Unit of Measurement of the eligible quantity, (2) by a monetary amount per unit of the Product or Service Instance in the given Unit of Measurement specified as a literal value of type float in combination with a currency, and (3) whether this prices includes local sales taxes, namely VAT. (Example: The price, including VAT, for 1 kg of a given material is 5 Euros per kg for 0 – 5 kg and 4 Euros for quantities above 5 kg). We do not model the full detail of local taxes.

Variants of Price Specification are Delivery Charge Specification (a conceptual entity that specifies the additional costs asked for delivery of a given Offering using a particular Delivery Method by the respective Business Entity) and Payment Charge Specification (a conceptual entity that specifies the additional costs asked for settling the payment after accepting a given Offering using a particular Payment Method). Both are defined in the GoodRelations ontology but not detailed in this paper for lack of space. Due to the complexity of pricing scenarios in various industries, it may be necessary to create extensions of this fundamental model of Price Specifications. Such can be done easily by importing and refining the GoodRelations ontology.

Validity: The time interval that specifies the period during which the offer is valid.

Currency: The Unit of Measurement for monetary values. For currencies, we suggest using the well-established ISO 4217 standard [15], which is the recommended encoding for currencies in international payment transactions.

Warranty Promise and Warranty Scope: The duration and scope of services that will be provided to a customer free of charge in case of a defect or malfunction.

Payment Method: A standardized procedure for transferring the monetary amount for a purchase.

Delivery Method: A standardized procedure for transferring the Product or Service Instance to the destination of fulfillment chosen by the customer.

Business Entity Type: The legal form, the size, the main line of business, the position in the value chain, or any combination thereof, of a Business Entity. From the ontological point of view, Business Entity Types are mostly roles that a Business Entity has in the market. Business Entity Types are important for specifying eligible customers, since Offerings are often meant only for Business Entities of a certain size, legal structure, or role in the value chain.

Country or Region: Countries or Regions are geographical or geopolitical areas. In GoodRelations, they are used for specifying the areas for which the Offering is valid.

Location of Sales or Service Provisioning: A location from which the specified Business Function on the particular Product or Service Instance is being offered by the Business Entity.

4 Ontology Coding

In the following, we motivate our choice of the ontology language and describe the coding of the ontology in OWL DLP. Note that the final ontology includes new conceptual entities for n-ary relations, since OWL supports binary relations only. The full ontology coding and description of elements is available at <http://www.heppnetz.de/projects/goodrelations/>.

4.1 Appropriate Ontology Language: RDFS, OWL DLP, or OWL DL?

Since GoodRelations ontology is meant to be of immediate practical value on the basis of currently available Semantic Web infrastructure, the choice of the appropriate ontology language is important. Eventually, we suggest using the OWL DL syntax for RDFS elements, i.e. a subset of the closure of OWL DLP, so that a lightweight RDFS-*style* reasoner can compute all practically relevant inferences while the ontology could also be used together with OWL DL ontologies and knowledge bases without making the resulting model become OWL Full. The language fragment we suggest is also, in practice, not touched by the layering problems between RDFS and OWL DL (for a discussion of these issues, see e.g. [16]). The GoodRelations ontology will only use the following language elements:

```
owl:Ontology, owl:Class, owl:ObjectProperty, owl:DatatypeProperty,
rdfs:subClassOf, rdfs:subPropertyOf, rdfs:comment, rdf:datatype, and
rdf:type
```

It also includes `rdfs:range` and `rdfs:domain`; however, this is meant only for helping tools generate forms when creating instance data. It is not necessary that the ontology repository computes the correct closure for the official RDFS semantics (but it would also not hurt). For a discussion of the semantics of domain and range in RDFS and OWL, see [17]. In combination with these domain and range statements, we use a few complex class definitions that define the union of multiple classes. However, again, this is used solely for specifying domains for object and datatype properties so that tools can generate forms easily for populating the ontology. For the correct interpretation of the offering data on a Web scale and query answering, it is not necessary that the ontology infrastructure computes the correct closure of this.

In short, the advantages of our choice are as follows:

- a) All data annotated using GoodRelations on the Web can be properly interpreted with just a RDFS-*style* reasoner, as long as this reasoner supports `owl:Class`, `owl:ObjectProperty`, and `owl:DatatypeProperty`. This is insofar important, as SPARQL queries on everything beyond a subset of RDFS are difficult and subject to ongoing research.
- b) At the same time, the ontology plus data in combination with OWL products and services ontologies will stay within OWL DL and can be used with a DL (or even DLP) reasoner – i.e., we do not end up in OWL Full.
- c) Also, we could otherwise not define important top-level concepts for product and services ontologies in GoodRelations, since they need to be imported by future releases of eClassOWL and other ontologies. If GoodRelations was based on RDFS elements like `rdfs:Class` and `rdfs:Property` for the definition of

its elements, importing the ontology and refining its elements in eClassOWL would turn eClassOWL into an OWL Full model if the exact semantics of the RDFS constructs was expected.

Our approach is in line with the proposal by Polleres et al. of using some restricted sets of RDF(S) inference with an extended version of SPARQL [18]. This would provide the foundation for reliable reasoning support for using (a future version of) SPARQL as a query language with GoodRelations-related data on a Web scale.

4.2 N-Ary Relations in OWL

We have seen in the domain capture that there exist several ternary and quaternary relations. Unfortunately, OWL supports only binary relationships in the form of object properties. This means that we have to find modeling work-arounds for the higher-arity relations in the domain. The basic pattern for handling this is (1) introducing an additional class for each of those relationships, of which instances act as placeholders that keep together all but one of the parameters of the n-ary relation and (2) defining a binary relation that links the remaining parameter with this placeholder instance. See [19] for more details.

4.3 Anonymous Instances and Existential Quantification

In the vast majority of e-commerce scenarios, the actual product or services individuals are not exposed on the Web. For example, if a Web shop says they will sell the Siemens cell phone model s1234 at \$ 100, there is mostly no information resource on the Web that describes this particular, tangible cell phone. From a logician's perspective, most offerings of commodities on the Web are just existentially quantifying the actual Product or Services Instances. The Web shop in the example actually says "There exists at least one thing X, for which holds ThisWebShop sells X and X instanceOf cellphone and X hasMakeAndModel Siemens-s1234". In such cases, there is no retrievable information or non-information resource reflecting the actual cell phone, on which ownership can be obtained, and which is gone once it has been sold. (Of course, there are a few situations where actual instances are being exposed on the Web. The most prominent ones are eBay auctions – there, a dedicated Web resource describing an actual product may exist. However, we can safely assume that in most commodity scenarios, the actual instances are only existentially quantified.)

From a theoretical perspective, this is no big deal. From a practical perspective, it is, for two reasons. First, existential quantification increases the computational cost of reasoning. In the series of popular ontology languages for the Semantic Web, only OWL DL and above and WSML DL support existential quantification. Second, and more importantly, the overhead for modeling Product or Services Properties on the basis of existing products and services ontologies, namely eClassOWL [7] is significant and modeling existential quantification in OWL is not at all straightforward for domain experts. For GoodRelations, we propose a very pragmatic work-around:

- a) We create two top-level ontology classes: (1) ActualProductOrService – an instance of this class is an actual Product or Service Instance. (2) ProductOrServicesSomeInstancesPlaceholder – an instance of this class is a placeholder for existentially quantified, anonymous Product or Service Instances.
- b) When we describe an actual Product or Service Instance (e.g. in an eBay offering), we make it an instance of the respective Product or Service Class (e.g. cell phone) *and* of the class ActualProductOrService. We can then describe this instance using all of the properties provided by products and services ontologies like eClassOWL.
- c) When we describe an anonymous Product or Service Instance (e.g. “we sell Siemens s1234 cell phones), we create an instance of the respective Product or Service Class (e.g. cell phone), and make it also an instance of the class ProductOrServicesSomeInstancesPlaceholder. We can then describe this instance using all of the properties provided by the products and services ontologies like eClassOWL.
- d) A query for a given type of products will return both the instances of ActualProductOrService and of ProductOrServicesSomeInstancesPlaceholder. Still, everybody can see quickly whether this instance is an actual cell phone or a placeholder, while the structure of queries regarding product features remains the same.

This approach works with very lightweight reasoners and is, more importantly, very well compatible with how current products and services ontologies are designed. One may object that the proxy instances are intangible objects and can thus not be instances of a class of tangible things like products. However, we think that the practical advantages of the work-around justify the small ontological inconsistency.

4.4 Ranges for Product and Service Attributes

Most quantitative properties of products or services are actually intervals and not single values. Even simple characteristics like the weight of a tangible object can in principle, only be specified as intervals (even if very small). Also, there are many scenarios in which queries for suitable products or services must allow ranges in the query and must reason properly about such ranges. For example, if someone is looking for a TV set with a screen size between 10 and 15 inches, a model with 12 inches must be reported as a match, and if someone has a piano that needs to be transported and weighs 120 kg, a company offering to transport all pianos up to 150 kg of weight must be found.

Now, there have been approaches of extending Web ontology languages, namely OWL, by support for datatype ranges. The most prominent approach is the OWL-Eu proposal by Pan and Horrocks [20]. This would allow proper reasoning about value ranges for OWL datatype properties. However, this extension is currently not widely supported by standard tooling; it has yet to find its way into mainstream Semantic Web infrastructure. Since GoodRelations aims at making Semantic Web-based E-Commerce a reality on the basis of current technology components, using OWL-Eu was not an option for the moment.

Instead, we use a pragmatic workaround that puts part of the work on the person expressing the query but at the same time requires only a standard RDFS or OWL reasoner that supports `rdfs:subPropertyOf`.

The approach is as follows: (1) We create an ontology class Quantitative Value. (2) All properties reflecting quantitative characteristics of products or services are represented as object properties with the range of “Quantitative Value”. (3) For each quantitative value, we create an instance of “Quantitative Value”. (4) We then attach the upper and lower limits of this value by two datatype properties (i.e. attributes) hasMinValue and hasMaxValue, and the unit of measurement by a datatype or object property hasUnitOfMeasurement. Now, querying for suitable object requires just specifying the lower and upper limits for the respective characteristics.

One may object that this means a lot of redundancy for those cases where the upper and the lower limit are practically the same. In order to mitigate this, we introduce a third datatype property hasValue, which is an `rdfs:subPropertyOf` both hasMinValue *and* of hasMaxValue. We can then simply say “myTVSet hasWeight X, X instance of QuantitativeValue, X hasValue 10, X hasUnitOfMeasurement “kg” “

As long as there is reasoner that computes that “X hasValue Y” implies “X hasMinValue Y, X hasMaxValue Y”, this shortcut will return the very same results.

5 Cool URIs for E-Commerce on the Semantic Web

When using GoodRelations, we need unique identifiers not only for the current Web resources, like corporate Web pages or pages in an e-shopping system, but for all conceptual elements as described in the domain capture in section 3. For example, we need unique identifiers for (1) business entities, (2) makes and models, (3) offerings, and (maybe) (4) available product instances. It is tempting to assume that we can simply take the URIs of existing Web resources for those purposes, but this is not advisable since the available Web resources often tangle multiple conceptual entities – the Web page describing a certain offer has one single URI, but it may contain the description of a product make and model, an offering, a price specification, a warranty promise, all at the same time – and, in the case of a small shop, even the business entity itself. When we move on to the Semantic Web, we need unique identifiers for every single entity of interest. In a nutshell, we can use the URIs of existing Web resources only as the identifiers for the Web resources themselves. For all other conceptual entities, we have to introduce new identifiers, unless such have already been defined for the Semantic Web elsewhere. For that, there exist three approaches:

- a) **Using blank RDF nodes** [23]; however, for all entities that may at any later point be linked to other entities this is not recommended, since it is then not possible to refer externally to such a node [24], which means that data from multiple sources cannot be merged easily.
- b) **Creating completely new URIs for all significant entities.** When doing so, one must have authority to define the meaning of the respective URI. For URIs in the HTTP scheme, this requires that one is the owner of the respective domain name space.

- c) **Creating hash URIs derived from the URI of the original Web resource.** In RDF, adding fragment identifiers to a URI creates a new identifier. It must be noted that also in this case, one must have authority to introduce new URIs in the respective domain space.

GoodRelations is agnostic of which approach is taken and makes no assumption on who defines the URI for a respective conceptual element; it is just necessary that distinct identifiers for distinct entities exist. Blank nodes are acceptable for entities that will never be linked to outside data (e.g. workarounds for n-ary relations, see section 4.2). The URI of the current, retrievable Web resource should become the URI of the conceptual entity Web Resource, and other conceptual entities, namely Offering, should point to this via `rdfs:seeAlso`. For a full discussion on the issues of naming entities on the Semantic Web, see [22, 24-26].

6 Evaluation

GoodRelations has been under development since 2005. During that time, several requirements from industrial partners were taken into account. While a full evaluation based on the annotation of large corporate data assets is still ongoing, there is already preliminary evidence that their key representational needs are now well covered by the ontology. Also, specific extensions for several very diverse domains have been created with ease, which indicates that the top-level structure seems to be appropriate. In the following, we show by means of examples of queries in SPARQL how the ontology can be used to query for respective product and services data on the Semantic Web. For the sample data used and screenshots of the results, please see our Technical report. We use the following prefixes:

```
PREFIX gr: <http://www.heppnetz.de/ontologies/goodrelations/v1#>
PREFIX ex: <http://www.heppnetz.de/ontologies/goodrelations/examples#>
```

Example 1: “Who sells cell phones and on which Web pages can I get more information on respective offerings?”

This transforms into the following SPARQL query:

```
SELECT ?business ?uri
WHERE {
  ?business gr:offers ?offering .
  ?offering gr:includesObject ?TypeAndQuantityNode .
  ?TypeAndQuantityNode gr:typeOfGood ?something .
  ?something rdf:type ex:Cellphone .
  ?offering gr:hasBusinessFunction gr:Sell .
  ?offering rdfs:seeAlso ?uri }
```

Example 2: Which offers of cell phones exists, what is the price, and where can I find the offering on the Web?

This transforms into the following SPARQL query:

```
SELECT ?offering ?uri ?price ?currency
WHERE {
  ?offering gr:includesObject ?TypeAndQuantityNode .
  ?TypeAndQuantityNode gr:typeOfGood ?something .
  ?something rdf:type ex:Cellphone .
```

```

?offering gr:hasBusinessFunction gr:Sell .
?offering rdfs:seeAlso ?uri .
?offering gr:hasPriceSpecification ?priceSpecification .
?priceSpecification rdf:type gr:UnitPriceSpecification .
?priceSpecification gr:hasCurrencyValue ?price .
?priceSpecification gr:hasCurrency ?currency . }

```

The result is as expected (note that offerings that do not have a price specification do not appear in here due to the structure of the query and the semantics of SPARQL).

Example 3: Who repairs at least one type of cell phone?

```

SELECT ?business ?businessuri ?offeringuri
WHERE { ?business gr:offers ?offering .
?offering gr:includesObject ?TypeAndQuantityNode .
?TypeAndQuantityNode gr:typeOfGood ?something .
?something rdf:type ex:Cellphone .
?offering gr:hasBusinessFunction gr:Repair .
?offering rdfs:seeAlso ?offeringuri .
?business rdfs:seeAlso ?businessuri . }

```

Note that this query does not guarantee that this shop will repair every make and model. We only know that it has offered to repair some objects of which is known that they are cell phones. Again, the result is as expected.

While a large-scale evaluation of the GoodRelations ontology is just starting, we can see that the ontology covers numerous typical usage scenarios. In particular, it is both generic and flexible, while at the same time requiring only minimal reasoning support.

7 Discussion

Related work to ours can be roughly grouped into three waves: The first wave: Gupta and Quasem [27] have stressed the potential that Semantic Web technology could reduce price dispersion in markets and thus make the market mechanism and e-commerce more efficient. Similarly, Fensel [28] described how ontologies can in principle support the integration of heterogeneous and distributed information in e-commerce, mainly based on catalogs of products, and which tasks are to be mastered. Obrst, Wray and Liu [2] discuss the main challenges of building and aligning ontologies for products and services in B2B e-commerce environments. Corcho and Gómez-Pérez [3] show how multiple standards for classifying products and services can be integrated using ontological mappings, and sketch a prototype implementation based on the WebODE platform. From that time there are a few very early transcripts of products and services classifications (UNSPSC and eCl@ss) into ontology languages [29, 30] and [31] available on the Web. The shortcomings of those early transcripts are described in [32].

The second wave: After 2003, several papers were published that analyze the technical problems of implementing the overall vision. Most of those works take into account the practical problems of product data management in real-world scenarios that the early papers did not, in particular regarding the amounts of data, the pace of evolution, the need to reuse existing standards, etc. Tolksdorf et al. [33] analyzed the potential impact of Semantic Web technology on B2C e-commerce and discussed the

technical and organizational challenges of the vision. They were, to our knowledge, the first to stress the lack of serious business ontologies, consensual identification schemes; open security, trust, and privacy issues, and requirements for the successful diffusion of the Semantic Web in the business world. Zhao and Sandahl [34] described the potential technical contributions and the required tools and structures for various processes of electronic procurement. Also, Zhao [35] and Zhao and Lövdahl [36] were the first to discuss the need for and difficulties of reusing existing standards when developing e-commerce ontologies. Di Noia et al. [37] have analyzed the complexity of matchmaking based on formal descriptions of offerings in electronic marketplaces. Beneventano [38] analyzed the mapping problems between eCI@ss and UNSPSC in detail and produced a prototype for managing respective mappings.

The third wave: Only recently, mature, practically usable ontologies for products and services have been derived from common classification standards, namely our ontology eClassOWL [7], as described in detail in [32]. The potential contribution of Semantic Web technology on future shopbots has been renewed in [39]. Most importantly, the enormous body of work done in Korea in the context of a large-scale research project materialized in the form of operational prototypes for ontology-based management of product data and recommender systems, as described in [4-6].

Apart from this latter work and some work on non-functional properties of Semantic Web services researchers, there is to our knowledge no mature proposal on how to encode commodity products and service offerings on the Semantic Web. The work done in Korea [4-6] is very impressive but is rather focused on B2B scenarios in closed settings. In particular, it relies on heavyweight formalisms. In a nutshell, GoodRelations is the first concrete technical proposal for implementing a widely shared vision.

Future directions: The following extensions and modifications of GoodRelations will be evaluated in the near future: (1) Microformat Variant and GRDDL Transformation and (2) **Proton Grounding** - Proton [40] is a popular general-purpose ontology and has some overlap with GoodRelations. We are evaluating the advantages and costs of grounding GoodRelations in Proton. (3) **More Advanced Price Modeling:** Several markets use more complex pricing schemas than those currently supported; see e.g. [14], BMEcat 1.2 [41], and BMEcat 2005 [42]. (4) **Resource Composition and Substitution:** A more sophisticated vocabulary for specifying resource composition and substitution knowledge would be often beneficial in the context of Semantic Web-based e-commerce. (5) **Multi-dependent Properties:** Some properties product instances or models are dependent on others. Currently, GoodRelations does not support modeling such properties. (6) **Integration with Catalog Data Standards for Harvesting Product Data:** For catalog data exchange on the basis of persistent publication on the Web, it may be beneficial to provide a Semantic Web vocabulary based on popular XML standards for B2B catalog data exchange, e.g. BMEcat 1.2 [41] and BMEcat 2005 [42]. We think that these should rather be ontologies in their own right, albeit aligned with GoodRelations.

8 Conclusion

We have (1) developed the representational requirements for an ontology that can be used for describing offerings of tangible goods and commodity services on the Web

and (2) coded a respective ontology using a small subset of current ontology languages, so that our ontology should scale very well on mainstream Semantic Web infrastructure available today. Our ontology is very flexible, while moderate in size, and supports value intervals plus existential quantification while posing only minimal requirements on the reasoning support of the ontology management infrastructure. Also, it should be compatible with some pragmatic reasoning support for SPARQL. We are currently working on tools that simplify the adoption of GoodRelations, e.g. on-line services that guide shop owners through the process of creating respective annotation data, and transformations from popular XML schemas for B2B catalog data interchange (namely BMEcat).

Acknowledgments: The author would like to thank (in alphabetical order) Jos de Bruijn, Doug Foxvog, Jacek Kopecky, Markus Linder, Axel Polleres, Andreas Radinger, Martin Schliefnig, and Katharina Siorpaes for numerous discussions and their invaluable feedback throughout three years of work in progress. The work presented in this paper has been supported by the Austrian BMVIT/FFG under the FIT-IT Semantic Systems project myOntology (grant no. 812515/9284), by a Young Researcher's Grant (Nachwuchsförderung 2005-2006) from the Leopold-Franzens-Universität Innsbruck, and by the European Commission under the project SUPER (FP6-026850).

References

- [1] Fensel, D., McGuinness, D.L., Schulten, E., Ng, W.K., Lim, E.-P., Yan, G.: Ontologies and Electronic Commerce. *IEEE Intelligent Systems* 16, 8–14 (2001)
- [2] Obrst, L., Wray, R.E., Liu, H.: Ontological Engineering for B2B E-Commerce. In: *Proceedings of the International Conference on Formal Ontology in Information Systems (FOIS 2001)*, Ogunquit, Maine, USA, October 17-19 (2001)
- [3] Corcho, O., Gómez-Pérez, A.: Solving Integration Problems of E-commerce Standards and Initiatives through Ontological Mappings. In: *Proceedings of the Workshop on E-Business and Intelligent Web at the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI 2001)*, Seattle, USA, August 5 (2001)
- [4] Lee, T., Chun, J., Shim, J., Lee, S.-g.: An Ontology-Based Product Recommender System for B2B Market places. *International Journal of Electronic Commerce* 11, 125–154 (2006)
- [5] Lee, H., Shim, J.: Conceptual Modeling of Product Information in e-Commerce. In: *Proceedings of the 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)*, Melbourne, Australia, July 11-13 (2007)
- [6] Lee, T., Lee, I.-h., Lee, S., Lee, S.-g., Kim, D., Chun, J., Lee, H., Shim, J.: Building an operational product ontology system. *Electronic Commerce Research and Applications* 5, 16–28 (2006)
- [7] Hepp, M.: eClassOWL. The Products and Services Ontology (retrieved May 20, 2008), <http://www.heppnetz.de/eclassowl/>
- [8] Uschold, M., Grüninger, M.: Ontologies: Principles, Methods, and Applications. *Knowledge Engineering Review* 11, 93–155 (1996)
- [9] United Nations Development Programme, Business Function Identifiers (BFI), <http://www.un-spssc.org/AdminFolder/documents/BFI.doc> (now offline)

- [10] United Nations Economic Commission for Europe (UN/CEFACT), Recommendation No. 20: Codes for Units of Measure Used in International Trade, CEFACT/ICG/2006/IC001 ed: UN/CEFACT Information Content Management Group (2006)
- [11] Milgate, M.: Goods and Commodities. In: Eatwell, J., Milgate, M., Newman, P. (eds.) *The New Palgrave: A Dictionary of Economics*, vol. 2, pp. 546–548. Macmillan and Stockton, London, New York (1987)
- [12] Guarino, N., Welty, C.A.: Evaluating Ontological Decisions with OntoClean. *Communications of the ACM* 45, 61–65 (2002)
- [13] Guarino, N., Welty, C.A.: An Overview of OntoClean. In: Staab, S., Studer, R. (eds.) *The Handbook on Ontologies*, pp. 151–172. Springer, Berlin (2004)
- [14] Kelkar, O., Leukel, J., Schmitz, V.: Price Modeling in Standards for Electronic Product Catalogs Based on XML. In: *Proceedings of the 11th International World Wide Web Conference (WWW 2002)*, Honolulu, Hawaii, USA, May 7–11 (2002)
- [15] ISO, ISO 4217:2001: Codes for the representation of currencies and funds (2001)
- [16] Eiter, T., Ianni, G., Polleres, A., Schindlauer, R., Tompits, H.: Reasoning with Rules and Ontologies. In: *Proceedings of the Reasoning Web 2006 Summer School*, Lisbon, Portugal, September 4–8 (2006)
- [17] de Bruijn, J., Lara, R., Polleres, A., Fensel, D.: OWL DL vs. OWL Flight: Conceptual Modeling and Reasoning for the Semantic Web. In: *Proceedings of the 14th International World Wide Web Conference (WWW 2005)*, Chiba, Japan, May 10–14 (2005)
- [18] Polleres, A., Scharffe, F., Schindlauer, R.: SPARQL++ for Mapping between RDF Vocabularies. In: Meersman, R., Tari, Z. (eds.) *OTM 2007, Part I. LNCS*, vol. 4803, pp. 878–896. Springer, Heidelberg (2007)
- [19] Noy, N., Rector, A., Hayes, P., Welty, C.: Defining N-ary Relations on the Semantic Web. W3C Working Group Note (April 12, 2006) (retrieved December 13, 2007), <http://www.w3.org/TR/swbp-n-aryRelations>
- [20] Pan, J.Z., Horrocks, I.: OWL-Eu: Adding Customised Datatypes into OWL. In: Gómez-Pérez, A., Euzenat, J. (eds.) *ESWC 2005. LNCS*, vol. 3532, pp. 153–166. Springer, Heidelberg (2005)
- [21] Booth, D.: Four Uses of a URL: Name, Concept, Web Location and Document Instance (retrieved January 15, 2007), http://www.w3.org/2002/11/dbooth-names/dbooth-names_clean.htm
- [22] Sauermann, L., Cyganiak, R., Völkel, M.: Cool URIs for the Semantic Web, German Research Center for Artificial Intelligence TM-07-01 (February 2007)
- [23] W3C, RDF Primer. W3C Recommendation (February 10, 2004) (retrieved December 10, 2007), <http://www.w3.org/TR/rdf-primer/>
- [24] Bizer, C., Cyganiak, R., Heath, T.: How to Publish Linked Data on the Web (retrieved July 25, 2007), <http://sites.wiwiss.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/>
- [25] Miles, A., Baker, T., Swick, R.: Best Practice Recipes for Publishing RDF Vocabularies (retrieved July 25, 2007), <http://www.w3.org/TR/swbp-vocab-pub/>
- [26] Lewis, R.: Dereferencing HTTP URIs. Draft Tag Finding (May 31, 2007) (retrieved July 25, 2007), <http://www.w3.org/2001/tag/doc/httpRange-14/2007-05-31/HttpRange-14.html>
- [27] Gupta, T., Qasem, A.: Reduction of price dispersion through Semantic E-commerce: A Position Paper. In: *Proceedings of the Semantic Web Workshop 2002*, Hawaii, USA, May 7 (2002)

- [28] Fensel, D., Ding, Y., Omelayenko, B., Schulten, E., Botquin, G., Brown, M., Flett, A.: Product Data Integration in B2B E-Commerce. *IEEE Intelligent Systems* 16, 54–59 (2001)
- [29] McGuinness, D.L.: UNSPSC Ontology in DAML+OIL (retrieved June 30, 2008), <http://www.ksl.stanford.edu/projects/DAML/UNSPSC.daml>
- [30] Klein, M.: DAML+OIL and RDF Schema representation of UNSPSC (retrieved June 30, 2008), <http://www.cs.vu.nl/~mca Klein/unspsc/>
- [31] Bizer, C., Wolk, J.: RDF Version of the eClass 4.1 Product Classification Schema (retrieved June 30, 2008), <http://www.wiwiss.fu-berlin.de/suhl/bizer/ecommerce/eClass-4.1.rdf>
- [32] Hepp, M.: Products and Services Ontologies: A Methodology for Deriving OWL Ontologies from Industrial Categorization Standards. *Int'l Journal on Semantic Web and Information Systems (IJSWIS)* 2, 72–99 (2006)
- [33] Tolksdorf, R., Bizer, C., Eckstein, R., Heese, R.: Business to Consumer Markets on the Semantic Web. In: *Proceedings of the On The Move to Meaningful Internet Systems OTM 2003 Workshops*, Catania, Sicily, Italy, November 3-7 (2003)
- [34] Zhao, Y., Sandahl, K.: Potential Advantages of Semantic Web for Internet Commerce. In: *Proceedings of the International Conference on Enterprise Information Systems (ICEIS)*, Angers, France, April 23-26 (2003)
- [35] Zhao, Y.: Develop the Ontology for Internet Commerce by Reusing Existing Standards. In: *Proceedings of the International Workshop on Semantic Web Foundations and Application Technologies (SWFAT)*, Nara, Japan, March 12 (2003)
- [36] Zhao, Y., Lövdahl, J.: A Reuse-Based Method of Developing the Ontology for E-Procurement. In: *Proceedings of the Nordic Conference on Web Services (NCWS)*, Växjö, Sweden, November 20-21 (2003)
- [37] Di Noia, T., Di Sciascio, E., Donini, F.M., Mongiello, M.: A System for Principled Matchmaking in an Electronic Marketplace. In: *Proceedings of the Twelfth International World Wide Web Conference (WWW 2003)*, Budapest, Hungary, May 20-24 (2003)
- [38] Beneventano, D., Guerra, F., Magnani, S., Vincini, M.: A Web Service based framework for the semantic mapping amongst product classification. *Journal of Electronic Commerce Research* 5, 114–127 (2004)
- [39] Fasli, M.: Shopbots: A Syntactic Present, A Semantic Future. *IEEE Internet Computing* 10, 69–75 (2006)
- [40] SEKT Consortium, PROTON Ontology (retrieved January 15, 2007), <http://proton.semanticweb.org/>
- [41] Schmitz, V., Kelkar, O., Pastoors, T., Renner, T., Hümpel, C.: *Specification BMEcat Version 1.2*, Stuttgart/Essen (2001)
- [42] Schmitz, V., Leukel, J., Kelkar, O.: *Specification BMEcat 2005*, Stuttgart/Essen (2005)