Ontology of units of measure

H. Rijgersberg[1]*, J.L. Top[12], M. Wigham[1]
[1]Agrotechnology & Food Sciences Group, Wageningen University and Research Centre,
P.O. Box 17, NL-6700 AA Wageningen, The Netherlands
{Hajo.Rijgersberg, Jan.Top, Mari.Wigham}@wur.nl
http://www.agrotechnologyandfood.wur.nl
[2]Vrije Universiteit Amsterdam, Faculty of Sciences, Computer Science,
De Boelelaan 1081A, NL-1081 HV Amsterdam, The Netherlands
http://www.few.vu.nl

Abstract

An ontology of units of measure is an important prerequisite for unambiguously exchanging and processing quantitative information. To evaluate existing ontologies of units, we compare the ontologies with a semi-formal description of the domain of units of measure, which we draft from textual descriptions of standards in the field. An important result of the analysis is that the ontologies only define subsets of the necessary concepts and relations identified in our reference description. On the basis of the description and the corresponding parts of the analyzed ontologies, we build a new ontology, called OUM (Ontology of Units of Measure and related concepts). Finally, we report on the ontology's application in web services and workflows (web applications and an add-in for Excel), leading to useful user applications.

1 Introduction

Formalization of units of measure and related concepts, such as quantities and dimensions, is important in exchanging and processing quantitative information. Many activities in different fields – not limited to the exact sciences only – depend heavily on unambiguous communication and interpretation of quantitative models and data. Standardized concepts allow scientists to formulate theories and to have their experiments reproduced. They also make reliable engineering possible. Currently, most of the contextual information needed to interpret mathematical and numerical information remains at the level of informal comments. As a consequence, this contextual information is often ambiguous and incomplete. For example, units of measure are frequently omitted when presenting scientific models, making the assumption that a default choice is shared by all readers. However, many scientists and engineers will agree that incomplete specification in the work of others is a major source of confusion and errors. This becomes even more manifest when models and data are processed automatically by software tools. Currently, as part of e-science and Semantic Web activities, vocabularies for computers are created (Hey & Trefethen, 2005). This supplements past practice when most emphasis in automating scientific computations was on numerical processing and visualization.

A common way to specify a computer vocabulary in recent years is using ontologies. The importance of an ontology of units of measure is recognized by the W3C Semantic Web Best Practices and Development (SWBPD) working group (W3C, 2004a). Such ontologies do exist, but they are not widely used yet, which may be related to the quality of the ontologies, which varies considerably. In order to evaluate the ontologies more objectively, we use a frame of reference, a semi-formal description of the domain of units of measure, which we draft from textual descriptions of standards in the field. On the basis of the description and the

---

*Corresponding author. E-mail address: Hajo.Rijgersberg@wur.nl

corresponding parts in the analyzed ontologies, we construct a new ontology of units of measure. Finally, we report on the ontology's application in web services and workflows (web applications and an add-in for Excel), leading to useful user applications.

2 A semi-formal description of the domain of units of measure

We selected the following sources as original and official references describing the domain of units and quantities, to distil our reference description from:

- Cohen, E.R., Giacomo, P.: Symbols, Units, Nomenclature and Fundamental Constants (1987),
- The CRC Handbook of Chemistry and Physics (1976),
- Taylor, B.N.: Guide for the use of the International System of Units (1995),
- The NIST Reference on Constants, Units, and Uncertainty (2004).

The selection is motivated as follows. The work of Cohen and Giacomo was compiled by the Commission for Symbols, Units, Nomenclature, Atomic Masses and Fundamental Constants (SUNAMCO commission) of the International Union of Pure and Applied Physics (IUPAP) and has been approved by the successive General Assemblies of the IUPAP held from 1948 to 1984. The CRC Handbook of Chemistry and Physics is a standard work which, among many other things, provides a detailed description of special systems of units used in electricity and magnetism, such as the cgs systems of units. This description is additional to Cohen & Giacomo (1987). It reflects definitions that were set by the S.U.N. commission (Symbols, Units and Nomenclature), predecessor of the above-mentioned SUNAMCO commission. Taylor (1995) is a guide for the use of the SI standard in the U.S. prepared by the National Institute of Standards and Technology (NIST). The document reflects the SI standard as described in the official ISO documents. It discusses fundamental aspects of the SI standard including classes of units of measure and the SI prefixes that are used to form decimal multiples and submultiples of units. NIST has also produced the NIST Reference on Constants, Units, and Uncertainty (2004) which describes, among other things, prefixes for binary multiples of units (units that are or should be used in information technology).

Based on the text sources we formulated a number of propositions that describe the domain of units of measure (see Table 1). The main concepts are:

- unit of measure
- prefix
- quantity
- measurement scale
- measure
- system of units
- dimension

These concepts all relate to enabling the expression of studied quantities in terms of standard quantities. For example, the length of a table can be expressed in terms of the length of the path traveled by light in vacuum during a time interval of 1/299 792 458 of a second, a standard quantity defining the metre.

Measurement scales usually have a number of categories or points referring to standard quantities. For example, the points of the Kelvin scale refer to triple points of metals or fluids under standardized conditions. An important aspect is that most units and scales refer to standard quantities *indirectly*. Usually they are defined in terms of *other* units of measure and scales, often using measures, which combine numerical values with units of measure or measurement scales. In this way, for example, the inch is defined in terms of the metre (“0.0254 m”).

Quantities combine metrological concepts with real-world phenomena. For example, the diameter of a steel cylinder relates a diameter (a metrological concept) to a steel cylinder (a real-world phenomenon). Quantities are classified according to their metrological concept. So the diameter of a steel cylinder is classified as a *diameter* rather than a *cylinder quantity*. Each class of quantities is expressed by a subset of units of measure or measurement scales. For example, length quantities are expressed using metre, inch, and so on.

Different kinds of units of measure exist: multiples and submultiples of units, compound units, and what we propose to call singular units. Multiples and submultiples of units combine a prefix and a unit. Examples of multiples and submultiples are kilogram and millisecond. Compound units are compositions of units using the mathematical operations multiplication, division or exponentiation[1]. Examples of compound units are cubic metre (m^3), pascal second (Pa·s), and candela per square centimetre (cd/cm^2). Units that are used as the elementary building blocks in forming multiples and submultiples of units, such as metre and pascal (units that have a special name), are not regarded as special in the standard literature sources. We argue, however, that they should be distinguished for the reason that only these units can be used to form the multiples and submultiples of units. We propose to use the term *singular unit* to denote these units.

Units of measure and quantities have a dimension. Dimensions are abstract properties of units and quantities neglecting their vectorial or tensorial character and all numerical factors including their sign. Dimensions can be expressed as the products of powers of base quantities of a system of units. For example, the mass dimension has an expressoipm of L=0, M=1, T=0, and so on in the SI, and L=–1, F=1, T=2 in the British system of units.

In order to achieve a coherent, interdependent set of units of measure in the wide variety of units that exist, they are organized in systems of units. The most widely used system of units is the International System of Units (SI). Other important systems of units are the British system and several cgs (centimetre gram second) systems, such as the Gaussian system of units. A system of units is based on a set of units chosen by convention to be the system’s base units, units that are considered to be mutually independent (i.e., can not be expressed in terms of each other). The units of measure of *derived* quantities – quantities defined in terms of the system’s base quantities – are expressed as products of powers of the base units.

Units of measure and quantities are commonly grouped in practice according to their use in a certain domain. For instance, the units newton, kilogram and metre per second squared, and the quantities force, mass and acceleration are grouped together according to their use in the mechanical domain. Specific units of measure and quantities may occur in more than one specific domain. For the purpose of grouping units of measure and quantities for practical use we propose to use an additional concept *application area*. We propose to define this concept on the basis of the fourteen categories distinguished in Cohen & Giacomo (1987) among which are mechanics, thermodynamics, and electricity and magnetism.

Table 1. Propositions describing the domain of units of measure, drafted from official text sources.

1. Units of measure, measurement scales, and measures express the extent of quantities.
2. A quantity can be expressed by one or more units of measure or measurement scales.
3. A unit of measure or measurement scale can be used for expressing one or more classes of quantities.
4. Units of measure are direct or indirect references to specific (standard and constant)

---

[1] Compound units must not be confused with derived units. The term *derived unit* only signifies the role of unit in a system of units, in contrast to its base units.

quantities.

5. Multiples and submultiples of units are a special kind of unit of measure.
6. Multiples and submultiples of units combine a singular unit and a prefix.
7. Prefixes represent conversion factors.
8. SI prefixes and binary prefixes are different kinds of prefixes.
9. SI prefixes represent powers of ten.
10. Binary prefixes represent tenth powers of two.
11. Compound units are a special kind of unit of measure.
12. A compound unit is composed from other units of measure using mathematical operations (multiplication, division or exponentiation).
13. Only units of measure with a special name can be used in the construction of multiples and submultiples of units and compound units.
14. Units of measure with a special name are called singular units.
15. A quantity represents a metrological aspect of a studied object, system, situation, etc.
16. A quantity relates to a studied object, system, event, etc. (proposed to be called phenomenon).
17. Quantities are classified on the basis of similarity in metrological aspect rather than the phenomena they relate to.
18. Four types of measurement scales that exist are nominal scales, ordinal scales, interval scales, and ratio scales.
19. Nominal scales have categories.
20. Ordinal scales have categories in a certain order.
21. Interval scales and ratio scales have points, which are related to quantities or phenomena in the real world.
22. Ratio scales additionally have a true zero point, representing an absolute zero.
23. Interval and ratio scales are related to units of measure.
24. A measure combines a numerical value with a unit of measure or measurement scale.
25. Measures are used for expressing conversion rules between units of measure.
26. A system of units is based on a set of units of measure chosen by convention to be the units of measure of the system’s base quantities, quantities that are considered to be mutually independent.
27. Derived quantities are defined in terms of products of powers of the base quantities of a system of units.
28. Units of measure of derived quantities of a system of units are expressed as products of powers of base units of the system.
29. A system of units has base dimensions and derived dimensions, which can be determined from the dimensions of a system’s base quantities and derived quantities.
30. Units of measure and quantities have a dimension.
31. Dimensions are abstract properties of units and quantities neglecting their vectorial or tensorial character and all numerical factors including their sign.
32. Dimensions can be expressed as the products of powers of base quantities of a system of units.
33. For the purpose of grouping units of measure and quantities for practical use, an additional concept *application area* is defined.
34. This concept has at least the fourteen categories as distinguished in Cohen & Giacomo (1987).

3 Analyzing vocabularies of units of measure

We analyzed selected ontologies of units of measure using the semi-formal description given in Table 1 as a frame of reference. The method is referred to as “analysis of ontology”, which is a part of the ontology evaluation approach by Gómez-Pérez (2001). It proposes a number of criteria based on earlier ontology evaluations, from which we select the following criteria:

- Completeness of the modeled scope
- Quality of formal definitions
- Consistency (understandability and extensibility)
- Completeness in and clarity of the natural language documentation

Completeness of the modeled scope indicates to what extent the main concepts in our frame of reference are present in the examined ontologies. *Quality of formal definitions* expresses how close the descriptions are to the studied objects. *Understandability* and *extensibility* concern more basic issues such as consistent naming, systematic inclusion of instances, and so on – in other words, how *consistent* the examined ontologies are. *Completeness in the natural language documentation* concerns the quality of the natural language descriptions of the modeled concepts.

We selected the following ontologies for analysis:

- EngMath,
- SUMO,
- ScadaOnWeb,
- SWEET Unit,
- OpenMath units and dimension CD groups.

These ontologies are among the best known ontologies of units. EngMath is an ontology for mathematical modeling in engineering, designed in the early 1990s. The ontology defines units, quantities, dimensions, and so on and was intended to be a foundation for other engineering ontologies (Gruber & Olsen, 1994). We examine the Ontolingua files as published in 1993 (EngMath, 1993). SUMO (Suggested Upper Merged Ontology) is the result of a collaborative effort involving the work of many researchers as part of the IEEE SUO effort which contains a section on quantities and units of measure. We examine the ontology code as published in 2003 (SUMO, 2003). The ScadaOnWeb approach to quantities and scales is identical to that defined in ISO 15926-2, a standard that specifies a conceptual model for the representation of technical information about process plants (Leal & Schröder, 2002). We examine the OWL files published in 2003 (ScadaOnWeb, 2003). SWEET Unit is part of the Semantic Web for Earth and Environmental Terminology (SWEET) project of NASA which provides a semantic framework for earth science initiatives (SchemaWeb, 2006). We examine the OWL files from 2004 (SWEET Unit, 2004). The OpenMath units and dimension CD groups are part of OpenMath, a standard for the representation of mathematical objects, allowing them to be exchanged between computer programs (Davenport & Naylor, 2003). We examine the code as presented in 2003 (OpenMath, 2003).

An important result of the analysis is that the ontologies only define subsets of the main concepts and propositions as distinguished in the reference description (see Table 2). Furthermore, we observe a number of discrepancies between the description and the ontologies, in general relating to not (properly) distinguishing concepts (in particular unit and quantity, measure and quantity, and measurement scale and unit of measure), not referring to predefined concepts (in particular multiples and submultiples of units do not refer to predefined prefixes and singular units), and inconsistent naming and incompleteness in the natural language definitions in the ontologies.

Table 2. Support of the main concepts and relations in the reference description of the domain of units by the selected ontologies.

Main concept or relation	Ontology				
	EngMath	SUMO	ScadaOnWeb	SWEET Unit	OpenMath
Unit of measure	√	√	-	√	√
Prefix	-	√	-	√	√
Quantity	√	√	√	-	-
Measurement scale	-	-	√	-	-
Measure	-	-	√	-	-
System of units	√	-		-	-
Dimension	√	-	√	-	√
Quantities formally refer to the units of measure that can be used for expressing them	√ (units refer to dimensions)	√	√ (quantities refer to measurement scales)	-	√ (units refer to dimensions)
Units of measure have formal definitions in terms of other units of measure or standard quantities	√	√	√	√	√
Multiples and submultiples of units formally refer to predefined prefixes	-	-	-	√	-

4 Ontology of Units of Measure and related concepts (OUM)

The Ontology of Units of Measure and related concepts (OUM) is based on the semi-formal description given in Table 1 and the corresponding parts in the analyzed ontologies (some of them given in Table 2). Figure 5 shows OUM's structure. Figures 2-4 show class diagrams of some of the ontology's classes and properties. OUM is modeled in OWL[2].

The ontology contains the main concepts listed in Section 2. A quantity is related to possible units and measurement scales by its properties `unit_of_measure` and `measurement_scale`. Multiples and submultiples of units refer to predefined prefixes using the property `prefix`. We define the class `Prefix` with property `factor` in order to represent the numerical factor of a prefix. Two subclasses are defined here: `SI_prefix` and `Binary_prefix`.

Units of measure and the points and categories of measurement scales have an explicit definition in terms of other units of measure, points or categories via the property `definition`. The value of a `definition` property is usually a measure, prescribing a conversion rule between the particular units. Ultimately, its range is `Quantity`, referring to standard quantities at the end of the definition chain.

`Quantity` has a property `phenomenon` of the type `Thing` to express its relation to a real-world object. `Quantity` has a range of subclasses such as `Length`, `Mass`, and `Time` to specify metrological aspects.

---

[2] The ontology can be freely downloaded from http://www.atoapps.nl/foodinformatics, Sec. "News".

Three subclasses for `Compound_unit` are defined (`Unit_division`, `Unit_multiplication`, and `Unit_exponentiation`) to represent the (nested) parts of compound units. We define the class `Measure` with properties `numerical_value` (range `Float`) and `unit_of_measure_or_measurement_scale` (range `Unit_of_measure` and `Measurement_scale`).

We take a pragmatic approach in modeling dimensions. Each of the combinations of base quantities and exponents is represented as a separate property, such as `SI_length_exponent`, `SI_mass_exponent`, `British_system_of_units_length_exponent`, and so on.

`System_of_units` has the properties `base_unit`, `derived_unit`, `base_quantity` and `derived_quantity`. We implement most of the prevailing systems of units as instances. We define the class `Application_area`, with the multiple valued properties `quantity`, `unit_of_measure` and `measurement_scale`. The fourteen categories from Cohen & Giacomo (1987) are defined as instances.

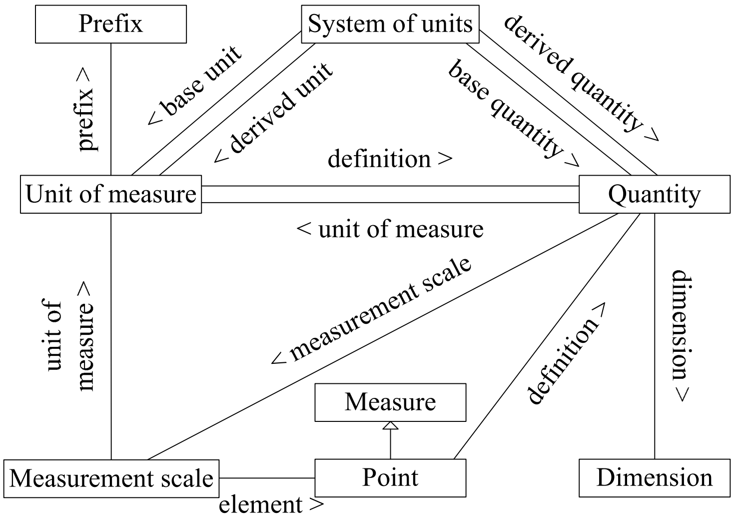


Figure 1. Simplified class diagram (UML) of OUM.

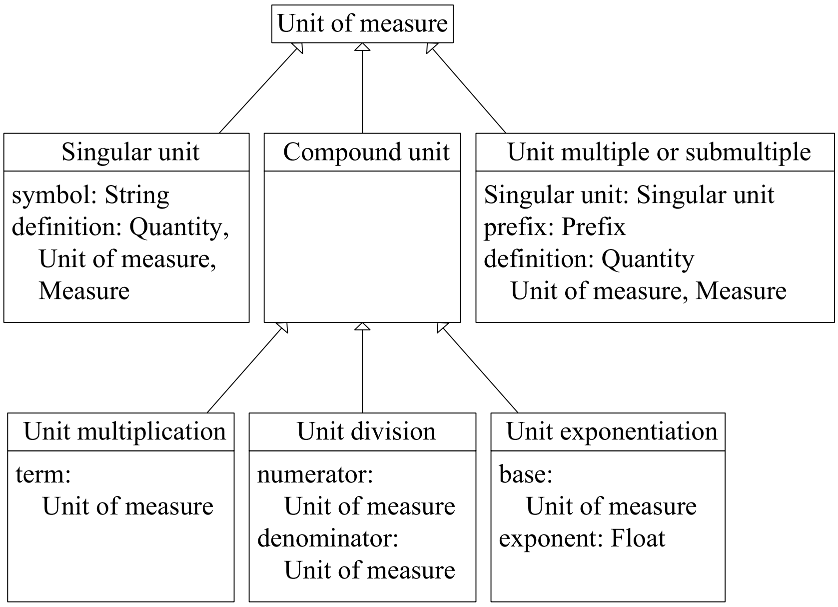


Figure 2. Class diagram (UML) of “Unit of measure” in OUM.

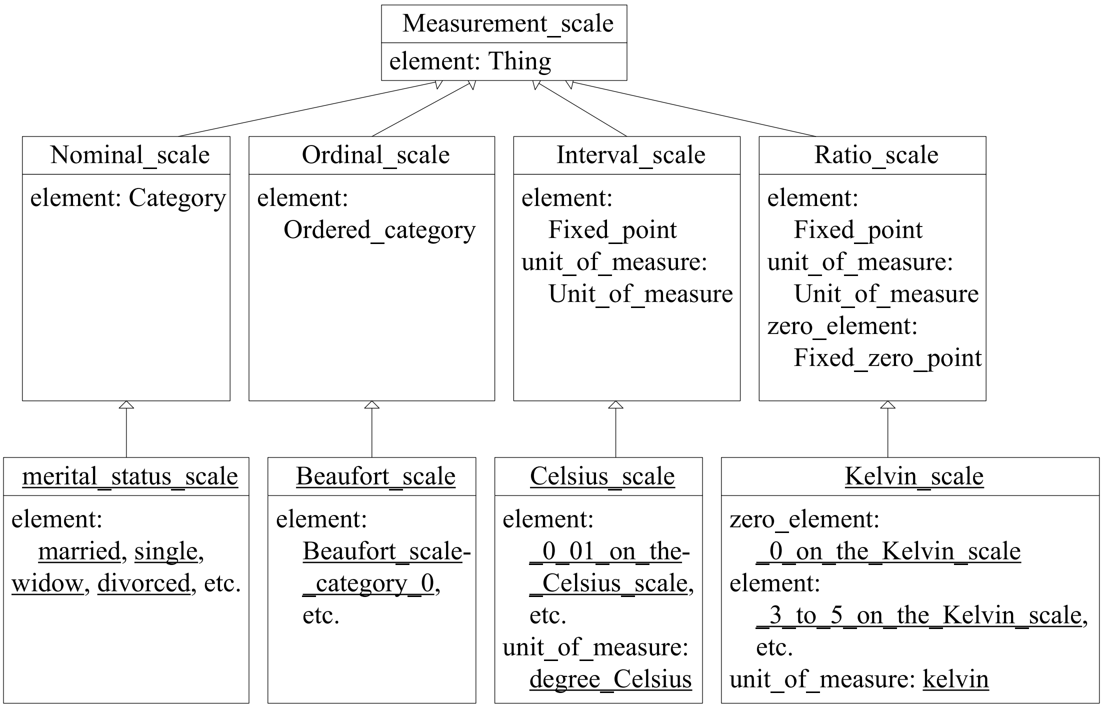


Figure 3. Class diagram (UML) of "Measurement scale" in OUM. Four instances of measurement scales are shown (underlined).

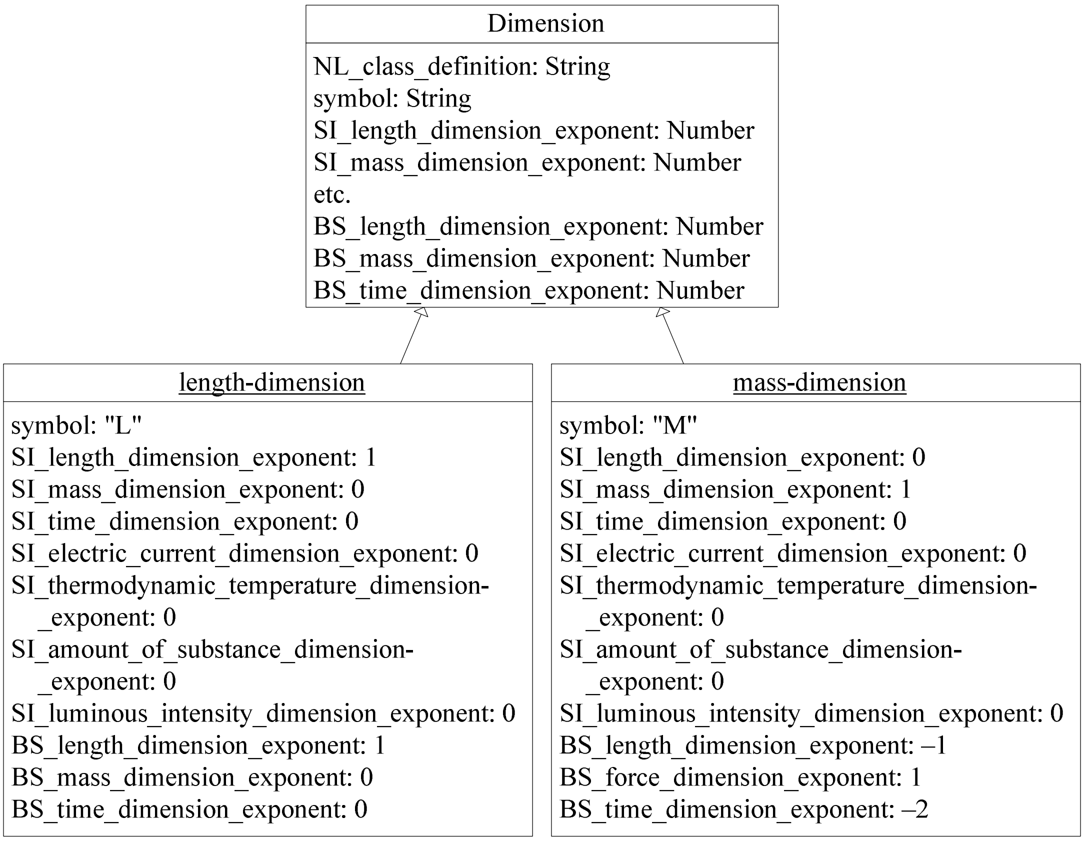


Figure 4. Class diagram (UML) of "Dimension" in OUM. Two instances of dimensions are shown (underlined). In the figure, the British system of units is abbreviated "BS".

5 Applying OUM

To demonstrate the usefulness of the OUM ontology, we have applied the vocabulary in web services and applications for the support of quantitative research processes:

- web services for extracting various types of information from the ontology and performing some basic functions using this information,
- web applications for combining the above-mentioned services to provide useful applications with user-friendly interfaces for the user,
- an Excel add-in with the same functionality,
- a spreadsheet enrichment tool which again uses the above-mentioned services.

The web services include functions such as retrieving possible units of measure for a given quantity, retrieving alternative units for a given unit of measure, etc. The services are implemented in Java and made available via a SOAP interface, so they can be used in any application regardless of the programming language or platform. The SOAP interface describes the necessary input parameters for the services, and what data the service returns.

We use the web services as components to build up complete user-friendly applications:

- finding symbols for a unit,
- finding symbols and units for a quantity,
- finding the conversion factor for two units,
- checking the consistency of an equation.

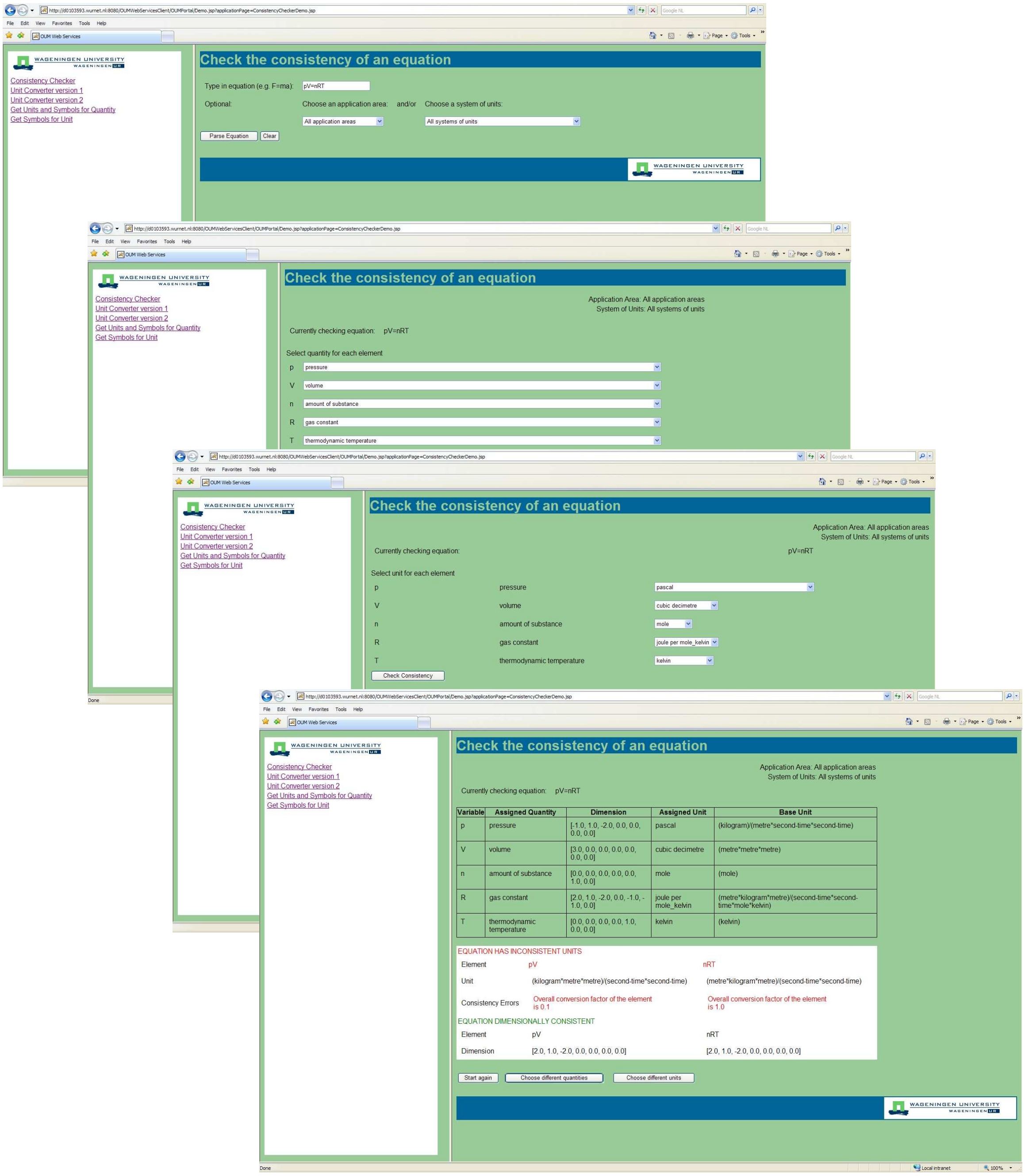
In “finding symbols for a unit”, based on a given unit the symbols for that unit are given. In “finding symbols and units for a quantity” the symbols and units are found for a quantity. In “finding the conversion factor for two units”, the conversion factor between two units is calculated on the basis of the definitions of the units in terms of a base SI unit. Finally, in “checking the consistency of an equation”, an equation is tested for unit and dimension consistency. The user enters a formula and then chooses quantities and units for the variables. The tool can then evaluate the consistency of this equation utilizing dimensional knowledge defined in OUM. Figure 6 shows screenshots of a test version of the web applications.

We evaluated the web applications with researchers. The researchers confirm the relevance and usefulness of such functions for their work. The format of a web application is however not so suitable. They would prefer that the tools be integrated in existing software such as Excel and Word. For this purpose we have applied the services in an add-in for Excel. The user selects a data block to perform a function such as unit conversion. Figure 7 shows the add-in in the form of a side panel.

Additionally, we have used the services in a spreadsheet enrichment tool. The aim of this tool is to enrich existing spreadsheets containing numerical data (legacy spreadsheets) with quantities, units of measure, measurement scales, etc. The tool interprets spreadsheets on the basis of string matching services, after which suggestions are proposed to the user. Also this tool supplies (simple) recognition of headers and cells. This tool is a .NET application. The tool is used in a research management system that we are currently developing, called Tiffany (Top & Broekstra, 2008).

Furthermore we have used OUM in a number of applications, from implementing workflows in Taverna, importing the ontology in domain ontologies and enriching numerical information, to performing conversion between quantities based on mathematical models from specific domains and unit conversion rules from the ontology.

The vocabulary enabled us to create the reported tools. Without the vocabulary most of the tools would have been difficult to develop. Now the underlying “database” was already available, and by updating the ontology the tools are immediately updated.



Check the consistency of an equation

Type in equation (e.g. F=ma): pV=nRT

Optional: Choose an application area: All application areas and/or Choose a system of units: All systems of units

Check the consistency of an equation

Application Area: All application areas
System of Units: All systems of units

Currently checking equation: pV=nRT

Select quantity for each element

p pressure
V volume
n amount of substance
R gas constant
T thermodynamic temperature

Check the consistency of an equation

Application Area: All application areas
System of Units: All systems of units

Currently checking equation: pV=nRT

Select unit for each element

p	pressure	pascal
V	volume	cubic decimetre
n	amount of substance	mole
R	gas constant	joule per mole_kelvin
T	thermodynamic temperature	kelvin

Check the consistency of an equation

Application Area: All application areas
System of Units: All systems of units

Currently checking equation: pV=nRT

Variable	Assigned Quantity	Dimension	Assigned Unit	Base Unit
p	pressure	[-1.0, 1.0, -2.0, 0.0, 0.0, 0.0, 0.0]	pascal	(kilogram)/(metre*second-time*second-time)
V	volume	[3.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]	cubic decimetre	(metre*metre*metre)
n	amount of substance	[0.0, 0.0, 0.0, 0.0, 0.0, 1.0, 0.0]	mole	(mole)
R	gas constant	[2.0, 1.0, -2.0, 0.0, -1.0, -1.0, 0.0]	joule per mole_kelvin	(metre*kilogram*metre)/(second-time*second-time*mole*kelvin)
T	thermodynamic temperature	[0.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0]	kelvin	(kelvin)

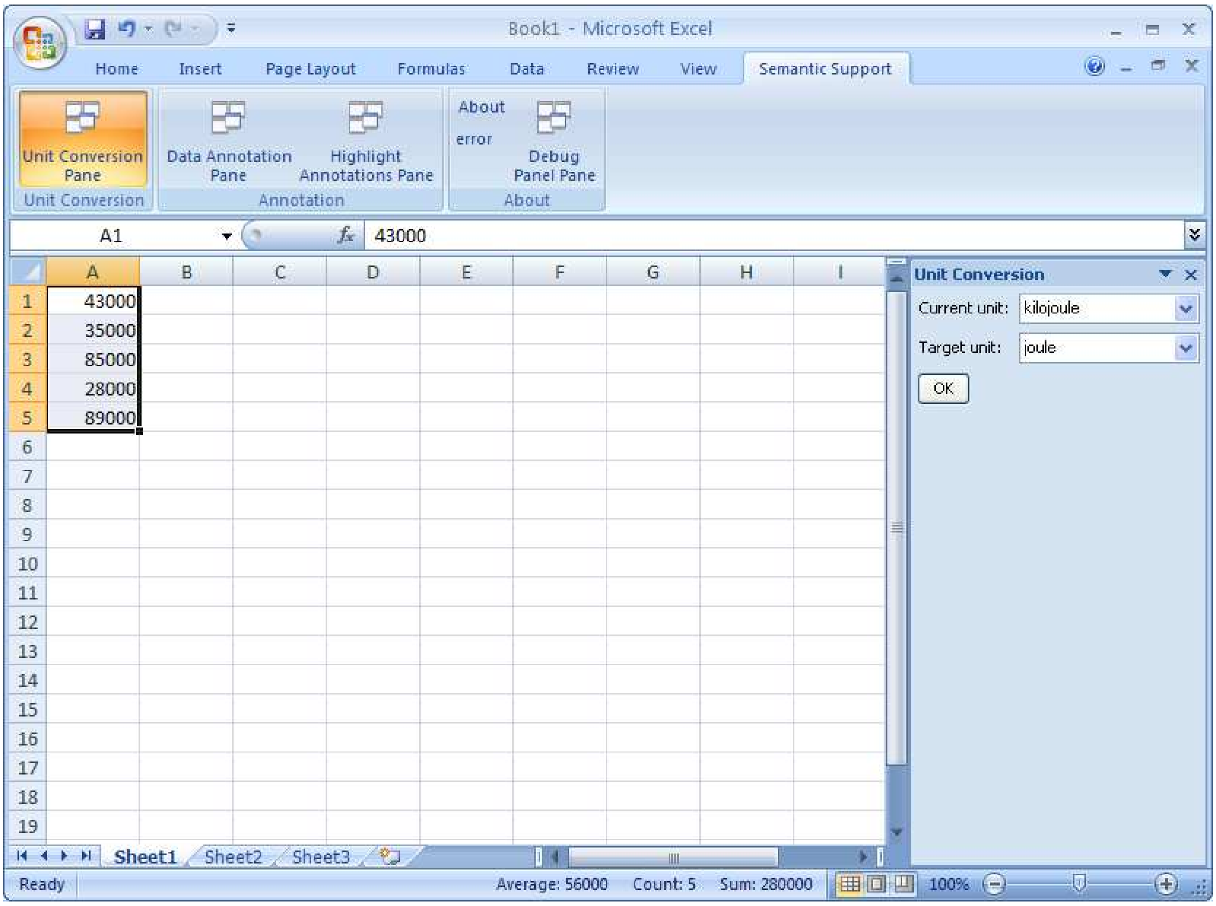
EQUATION HAS INCONSISTENT UNITS

Element	pV	nRT
Unit	(kilogram*metre*metre)/(second-time*second-time)	(metre*kilogram*metre)/(second-time*second-time)
Consistency Errors	Overall conversion factor of the element is 0.1	Overall conversion factor of the element is 1.0

EQUATION DIMENSIONALLY CONSISTENT

Element	pV	nRT
Dimension	[2.0, 1.0, -2.0, 0.0, 0.0, 0.0, 0.0]	[2.0, 1.0, -2.0, 0.0, 0.0, 0.0, 0.0]

Figure 6. Screenshots of the unit and dimension consistency checker web application.



A1 43000

	A
1	43000
2	35000
3	85000
4	28000
5	89000

Current unit: kilojoule

Target unit: joule

Average: 56000 Count: 5 Sum: 280000

Figure 7. Data conversion side panel add-in for Excel.

6 Discussion

It is quite surprising how intricate a seemingly simple framework of units of measure, quantities, dimensions, etc., can be. In our analysis of the original paper standards we encountered several pitfalls and peculiarities with respect to modeling these concepts. Interestingly, we did not run into many contradictions between the different paper sources. In fact they appeared to be highly complementary. However, a number of conceptual issues appeared to be more difficult to cope with. Here we discuss some of them.

In principle, one could decide to model metrological concepts separately from the notion of quantity. This way quantities can refer to both a metrological concept and a phenomenon, instead of that metrological concepts are inherent to quantity classes. For instance, length and mass could be defined as independent metrological classes, where quantities refer to e.g. “the length of my table” and “the mass of an electron”. In that case, the notion quantity boils down to a property (of “my table” and “an electron”) only. This seems to be an elegant alternative solution in many cases. The diameter of a sphere can then be said to have metrological concept “diameter” and phenomenon “a sphere”.

It is not straightforward to distinguish quantities from dimensions. In contrast to quantities, dimensions are dependent on the selected system of units. For example, a mass is a quantity. The expression of the mass dimension in SI (L=0, M=1, T=0, and so on) is different from how it is defined in the British system of units (L=–1, F=1, T=2), as mentioned earlier. In our approach we define system-of-units specific dimension exponent properties for dimensions, such as `SI_length_dimension_exponent`, `British_system_of_units_force_dimension_exponent`, etc.

The subtle distinction between a unit of measure and a measurement scale is not always properly recognized. This distinction is in particular relevant when considering the Celsius scale and its unit, the degree Celsius. 3 °C on the Celsius scale is something different than 3 °C in units of measure. The former indicates an absolute temperature equivalent to 276.15 K and the latter a temperature difference of 3 K. We have dealt with this issue by defining both units and measurement scales, which enabled us to define the degree Celsius (a unit) and the Celsius scale (a measurement scale). Measures can refer to a unit or a measurement scale.

We have defined many multiples and submultiples of units for the reason that one can better define too many than too few. Moreover, it is easier to define all combinations of prefixes to a particular singular unit than to attempt to make a selection on the basis of common usage.

How complete is the ontology? We can express this by indicating that the SI and several physical domains (from thermodynamics to quantum physics) are now present in the ontology. The ontology contains a set of specific length units from the typographical domain, illustrating that different units for specific (in this case: length) quantities can be defined. We have defined some phenomena in the ontology in order to be able to define base units of systems of units. For example, the metre is defined explicitly in terms of the length of the path traveled by light in vacuum during a time interval of 1/299 792 458 of a second in the ontology.

We use OWL (W3C 2004b) to specify the ontology. OWL is an emerging standard, designed by W3C (W3C, 2005) and can be used, like RDF, for specifying knowledge in the Semantic Web. The choice for OWL is motivated by it permitting us to make restrictions on property values, which we consider to be a required feature in the design of our ontology.

We have integrated certain services in Excel using an add-in. It is important to make this step towards data support in popular software since there is a high number of potential scientific users of this kind of functionality. The support is most convenient in two cases: if one has to work with units or concepts that one is not totally familiar with, or if certain actions (such as data conversion) often have to take place. Suggestions (for example for the source unit and target unit during conversion) that are made by the add-in will improve as the table format becomes more formal. In this case the headers will often be known. The headers may include useful information, for example, the units of the numerical data. So the user does not have to specify the source unit and, moreover, the target units can be restricted to compatible units.

In the field of unit conversion a lot of tools exist, many of them on-line, but these tools are not semantic – the underlying knowledge is not formal and open. Advantages of formal, open vocabulary are that software developers can share this information and the vocabulary can be updated at a central platform. Current unit converters do not distinguish the concept “quantity”. At most, units are grouped under headers that represent quantities in the UI. Unit and dimension consistency checkers do exist but limited to units, and with limited numbers of units. We blame this, once again, on the lack of an adequate vocabulary.

The analysis of existing ontologies was tough because it was difficult to find descriptive information of the ontologies. As a consequence, one has to inspect the ontologies. However this approach is full of pitfalls as it is no sinecure to load ontology code in browsers. Often one has to examine the ASCII code as a consequence. It is also difficult to contact the authors of the ontologies, something we have attempted in this work but succeeded in only partly.

7 Conclusion

In this paper we have drafted a semi-formal description of units of measure from textual sources, which we used to analyze existing ontologies of units of measure and to build a new ontology, preserving relevant ingredients from the existing ontologies. In general in the existing approaches, concepts are not (properly) distinguished from each other. Our goal was to remain close to the official documents.

We have written software that extracts various types of information from the ontology and performs some functions using this information. We have made this software available as web services, which can be called up via a web site and integrated in workflows and user applications. The knowledge in OUM is available via these web services, providing an easy interface to software developers. We have used the web services as components in workflows to build up complete user-friendly applications, such as a consistency checker for equations and a data conversion add-in for Excel.

We regard formalization of units of measure and related concepts as a first step towards formalization of quantitative information (data and mathematical models). Our next ambition is to formalize the structure of models and data as a means to represent the relations between quantities and their context and the underlying scientific reasoning process. This is crucial in interpreting and processing quantitative information automatically in the future.

Acknowledgements

We carried out this work within the context of the Virtual Laboratory for e-Science project (www.vl-e.nl), in the Food Informatics subprogram. This project was supported by a BSIK grant from the Dutch Ministry of Education, Culture and Science and was part of the Information and Communication Technologies innovation program of the Ministry of Economic Affairs. The Dutch Ministry of Agriculture provided additional funding. Thanks to Remko van Brakel, Jeen Broekstra, David Leal, Bill Naylor, and the anonymous reviewers of the article's earlier versions for valuable comments and reflections.

References

Cohen, E.R. and Giacomo, P. (1987). Symbols, units, nomenclature and fundamental constants in physics. 1987 revision. Document I.U.P.A.P.-25 (SUNAMCO 87-1). International Union of Pure and Applied Physics, SUNAMCO Commission.

Weast, R.C., ed. (1987). *Handbook of chemistry and physics.* Cleveland, Ohio: CRC Press.
Davenport, J.H. and Naylor, W.A. (2003). Units and dimensions in OpenMath. Available at: http://www.openmath.org/cocoon/openmath/documents/Units.pdf.

EngMath (1993). Available at: http://www.ksl.stanford.edu/knowledge-sharing/ontologies/html/standard-units/standard-units.lisp.html.

Gómez-Pérez, A. (2001). Evaluation of ontologies. *International Journal of Intelligent Systems*, 16, 391-409.

Gruber, T.R. and Olsen, G.R. (1994). An ontology for engineering mathematics. In Doyle, J., Torasso, P. and Sandewall, E. eds. Proceedings Fourth International Conference on Principles of Knowledge Representation and Reasoning. San Mateo, California: Morgan Kaufmann. Available at: http://www.ksl.stanford.edu/knowledge-sharing/ontologies/html/engineering-math.text.html.

Hey, T. and Trefethen, A. (2005). Cyberinfrastructure for e-Science. *Science,* 308, 817-821, doi:10.1126/science.1110410.

Leal, D. and Schröder, A. (2002). RDF vocabulary for physical properties, quantities and units. Available at: http://www.s-ten.eu/scadaonweb/NOTE-units/2002-08-05/NOTE-units.html.

OpenMath (2003). units_metric1, 3.0, 2003. Available at: http://www.openmath.org.

Rijgersberg, H. and Top, J. (2007). Ontology of Units of Measure and related concepts (OUM). Available at: http://www.atoapps.nl/foodinformatics, Sec. “News”.

SchemaWeb – SWEET Unit. (2006). Available at: http://www.schemaweb.info

ScadaOnWeb (2003). Available at: http://www.s-ten.eu.

SUMO (2003). Available at: http://www.ontologyportal.org.

SWEET Unit. (2004). Available at: http://sweet.jpl.nasa.gov/ontology/units.owl.

Taylor, B.N. (1995). *Guide for the use of the International System of Units (SI). 1995 edn.* U.S. Government Printing Office Washington, DC 20402.

The NIST Reference on Constants, Units, and Uncertainty. International System of Units (SI). Prefixes for binary multiples (2004). Available at: http://physics.nist.gov/cuu/Units/binary.html.

Top, J. and Broekstra, J. (2008). Tiffany: sharing and managing knowledge in food science. Keynote in ISMICK, Brasil.

W3C. (2004a). Semantic Web Best Practices and Deployment (SWBPD) Working Group Charter. World Wide Web Consortium (W3C). Available at: http://www.w3.org/2003/12/swa/swbpd-charter.

W3C. (2004b). Web Ontology Language (OWL). World Wide Web Consortium (W3C). Available at: http://www.w3.org/2004/OWL.

W3C. (2005). W3C World Wide Web Consortium. Leading the Web to Its Full Potential… World Wide Web Consortium (W3C). Available at: http://www.w3.org.