AMI Consult

OWMS Review

Final report

12 December 2008

AMI Consult S.àr.l.
Banque et Caisse d'Epargne de l'Etat IBAN LU15 0019 1055 9535 7000 – Fortis Bank Luxembourg IBAN LU58 0030 3181 5226 0000
Registre de Commerce Luxembourg: B 79.527 – Autorisation gouvernementale: 93529, 93529/A
Capital social: EUR 12.500 – Numéro matricule TVA: 2000 2429 174 – Numéro TVA Intracommunautaire: LU 18605040

Colophon

General

Client and policy accountability: Ministry of the Interior and Kingdom Relations (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties; http://www.minbzk.nl/)

Under supervision of: Overheid heeft *Antwoord*© (http://www.overheidheeftantwoord.nl/)

Report written by: Makx Dekkers, Thomas Baker and Mikael Nilsson

Rights

The Hague, December 2008

The Review Team:

Makx Dekkers has extensive experience in information technology, standardisation and international co-operation. The main areas of his activities are development, standardisation and application of solutions for information management and exchange with emphasis on interoperability from both a strategic as well as a technical perspective. He currently is the Managing Director of the Dublin Core Metadata Initiative. He is also involved in the Europeana project for Cultural Heritage and various activities related to eGovernment and public sector information. Previously, he worked in management positions at Pica, centre for library automation in the Netherlands and at the management consultancy department of PricewaterhouseCoopers in the Netherlands and Luxembourg. Makx has the nationality of the Netherlands and lives and works in Barcelona, Spain.

Thomas Baker is a member of the Directorate of the Dublin Core Metadata Initiative (DCMI), chair of the DCMI Usage Board and co-chair of the DCMI Architecture Forum. He lives in Germany, where he has worked as a researcher at National Research Center for Informatics and Mathematics (GMD), Fraunhofer-Gesellschaft, and the Göttingen State and University Library. He is currently co-chair of the application profile working group of Kompetenzzentrum Interoperable Metadaten (KIM) and the Semantic Web Deployment Working Group of the World-Wide Web Consortium.

Mikael Nilsson is a PhD student in the Knowledge Management Research Group at the Royal Institute of Technology in Stockholm and Uppsala University, and co-founder of the Semantic Web consulting firm Knowmania. He specializes in Metadata Interoperability, and currently co-chairs the DCMI Architecture Forum and the IEEE LTSC/DCMI Joint Task force. An active participant in DCMI, ISO/IEC JTC1 SC36 and IEEE LTSC, he is the co-author of multiple recent DCMI recommendations and other documentation, as well as upcoming IEEE metadata standards.

Table of Contents

Samenvatting

Achtergrond

Het document dat voor u ligt is het resultaat van een "collegiale kwaliteitsbeoordeling" van de "Overheid.nl Web Metadata Standaard" (OWMS) die op verzoek van het programma "Overheid heeft *Antwoord*©" in oktober en november 2008 uitgevoerd is door een team van deskundigen (Makx Dekkers, Thomas Baker en Mikael Nilsson).

Het onderzoek is uitgevoerd op basis van de volgende vragen:

1. Is OWMS semantisch in lijn met de semantiek van Dublin Core?
2. Is OWMS conceptueel in lijn met het Abstract Model van Dublin Core?
3. Is het OWMS technisch framework geschikt om zinvolle validatie van metadata mogelijk te maken?
4. Zijn relevante doorgroeimogelijkheden voor OWMS te zien?
5. Is OWMS als Europese overheidsstandaard toepasbaar?

Resultaten

Bij de beschouwing van bovenstaande vragen is het onderzoek tot de volgende conclusies gekomen:

Analyse semantiek

De semantiek van OWMS 3.5 is in grote lijnen in overeenstemming met de definities van de DCMI Terms. Aanvullende eigenschappen in de "mantel" zijn ook in lijn met de aanpak die ten grondslag ligt aan de DCMI Terms. Twee belangrijke problemen zijn vastgesteld: (a) het opnemen van contactinformatie in `dcterms:contributor` is in strijd met het een-op-een-principe, en (b) de definitie van het toepassing-specifieke element `overheidvac:revisionHistory` is niet in overeenstemming is met de rest van het model. Aangezien de evaluatie volgens nadere afspraak alleen betrekking heeft op de centrale componenten ("kern" en "mantel") is het laatste punt niet relevant.

Conformiteit Abstract Model

Het conceptuele model van OWMS is een verbeterde versie van het model dat wordt beschreven in de DCMI XML encoding guidelines van 2003 en niet gebaseerd op het actuele DCMI Abstract Model. In tegenstelling tot het DCMI Abstract Model worden in het OWMS-model de indicatie van syntaxcodeerschema's en taal niet direct aan de "value strings" gerelateerd zoals in het DCMI Abstract Model. Bovendien beschouwt het DCMI Abstract Model "taal" niet als een codeerschema. Hoewel het OWMS-model een onderscheid maakt tussen syntaxcodeerschema's en waardelijsten, wordt dit onderscheid in de syntax niet uitgedrukt (zie hieronder). Echter, de evaluatie concludeert dat er geen principiële belemmeringen zijn voor een conversie van metadata die gebaseerd is op het OWMS-model naar metadata volgens het DCMI Abstract Model, hoewel in specifieke gevallen een speciale behandeling van codeerschema's noodzakelijk kan zijn.

Analyse technisch framework

Het XML-framework is als adequaat beoordeeld, afgezien van het feit dat het framework geen complete implementatie van het conceptuele model is. In het bijzonder maakt de XML-codering geen onderscheid tussen syntaxcodeerschema's en

waardelijsten. De reviewers bevelen aan om ofwel het technisch framework aan te passen zodat het complete conceptuele model geïmplementeerd wordt, ofwel het conceptuele model aan te passen zodat het een afspiegeling is van het impliciete model dat aan de XML-codering ten grondslag ligt. De beschreven HTML-codering implementeert het model achter de XML-benadering, met uitzondering van de identificatie van de taal. Er worden geen directe acties aanbevolen op het gebied van nieuwe XML-ontwikkelingen.

Toekomstige ontwikkelingen

Hoewel OWMS in wezen een methode is om specifieke XML-talen te ontwikkelen, is het mogelijk om metadata in de centrale component van OWMS (de "kern" en de "mantel") naar RDF te transformeren met speciale behandeling van codeerschema's. "SKOSificatie" van de waardelijsten (waarbij aan de individuele termen van een lijst een URI toegekend wordt en de beschrijving van de termen in RDF uitgedrukt wordt volgens de SKOS-specificatie) zou het mogelijk maken dat andere organisaties de termen in de OWMS-waardelijsten kunnen gebruiken. Omgekeerd zou OWMS, als in de metadata naar termen in een waardelijst verwezen wordt met een URI, gebruik kunnen maken van waardelijsten die elders op een op SKOS gebaseerded wijze gedefiniëerd en onderhouden worden, bijvoorbeeld Eurovoc, de Europese thesaurus van administratieve terminologie, of GEMET, de Europese thesaurus van milieu-terminologie.

Europese toepasbaarheid

OWMS zou kunnen fungeren als een nuttig voorbeeld voor andere landen. Voor de uitwisseling van ervaringen met anderen is het wel nodig om informatie over OWMS in het Engels beschikbaar te maken. Door contacten te leggen met initiatieven in andere landen en door deel te nemen aan internationale platforms en conferenties kan het OWMS-team ervaringen met soortgelijke initiatieven uitwisselen.

Algemene conclusies

1. De semantische definities en het XML-framework van OWMS worden als adequaat beoordeeld.
2. Er is een probleem met de consistentie van de modellering, in de zin dat er een verschil bestaat tussen het conceptuele model, het impliciete model dat ten grondslag ligt aan de XML-codering en het model achter de HTML-specificatie. De reviewers raden aan dat er een keuze gemaakt wordt voor een consistent model, en dat de toepassing van dit model consequent beheerd wordt voor de centrale componenten van OWMS.
3. Er zijn geen fundamentele problemen om in de toekomst een meer semantische benadering te kiezen: conversie van metadata gemaakt op basis van de huidige model zou mogelijk zijn met een speciale behandeling van codeerschema's, hoewel daarbij moet worden opgemerkt dat de hoeveelheid hard te coderen conversieregels groeit naarmate er meer codeerschema's worden toegevoegd.
4. Voor de langere termijn kan een migratie naar een meer semantische benadering overwogen worden. Daarbij zouden waardelijsten en organisatienamen als aparte collecties (thesauri en name authorities) op basis van SKOS gedefiniëerd en beheerd kunnen worden. Verder zou bij verdere ontwikkeling van het model overwogen kunnen worden om het meer in lijn te brengen met het DCMI Abstract Model.

Executive Summary

Background and scope

The present document is the result of a "collegial quality assessment" of the "Overheid.nl Web Metadata Standaard" (OWMS) requested by the program "Overheid heeft Antwoord©" and conducted in October and November 2008 by a team of experts (Thomas Baker, Makx Dekkers and Mikael Nilsson).

The review was undertaken based on the following set of questions:

1. To what extent is OWMS semantically in line with Dublin Core semantics?
2. To what extent is OWMS conceptually in line with the DCMI Abstract Model?
3. To what extent is the technical framework suitable to enable a meaningful validation of instance metadata?
4. Which recommendations can be given from a Dublin Core perspective for further development of OWMS?
5. To what extent is the approach of OWMS more widely applicable, for example in a European context and the DC-Government Application Profile?

Results

The review looked at the above questions and came to the following conclusions:

Analysis of semantics

The semantics of OWMS 3.5 are by and large in line with the definitions of DCMI Terms. Additional properties of the "mantel" are also in line with the approach underlying the DCMI terms. Two major issues were identified: (a) the inclusion of contact information with `dcterms:contributor` which is a violation of the one-to-one principle, and (b) the definition of the application extension `overheidvac:revisionHistory` which is not in line with the rest of the model. This second point was declared out of scope for the review.

Conformity Abstract Model

The conceptual model of OWMS is an improved version of the model that is described in the DCMI XML encoding guidelines of 2003, rather than being based on the DCMI Abstract Model. One major difference with the DCMI Abstract Model is that OWMS does not attach syntax encoding schemes and languages directly to value strings the way the DCMI Abstract Model does. Moreover, the DCMI Abstract Model does not consider language to be a type of encoding scheme. The model differentiates between syntax encoding schemes and vocabulary encoding schemes in a manner consistent with the DCMI Abstract Model, though the distinction is not reflected in syntax (see below). However, the review concluded that conversion of instance metadata based on the OWMS model to metadata conforming to the DCMI Abstract Model would be fairly straightforward, although in specific cases there may be a need for special treatment of encoding schemes.

Analysis technical framework

The XML framework is considered to be adequate. However, it does not completely implement the conceptual model. Particularly, the XML encoding does not distinguish

between vocabulary and syntax encoding schemes. It is recommended to either modify the technical framework to reflect the conceptual model, or to adapt the conceptual model to reflect the implicit model underlying the XML approach. The HTML encoding implements the model behind the XML approach but does not allow the identification of the language of the described resource. No immediate actions are recommended regarding new XML developments.

Future developments

While OWMS is essentially a method for generating customized XML languages, the "Core Component" of OWMS (the "kern" and the "mantel") should be transformable into RDF with special treatment of encoding schemes. "SKOSification" (expression in RDF according to the SKOS specification with assignment of URIs to the individual terms in the vocabulary) of the "waardelijsten" would allow other organizations to use those vocabularies as vocabulary encoding schemes. Conversely, OWMS could adopt an approach whereby controlled vocabulary terms are referred to by URI, which would allow use of SKOSified versions of external vocabularies like Eurovoc, the European thesaurus of administrative terms, or GEMET, the European thesaurus of environmental terminology.

European applicability

OWMS could act as a useful example for other countries. For the exchange of experience with others it would be necessary to provide information about OWMS in English. Liaising with other initiatives and participating in international platforms and conferences could help the OWMS team to learn from experiences elsewhere.

Overall conclusions

1. The semantic definitions and the XML Framework that OWMS is based on are considered to be adequate.
2. There is a problem with the consistency of the modelling, as the conceptual model, the implicit model underlying the XML approach and the model behind the HTML encoding are slightly different. The reviewers recommend that a choice is made for one consistent model and that the application of that model be consistently enforced for the "Core Component" of OWMS.
3. There are no fundamental problems if at some point in the future a decision were to be taken to upgrade the OWMS approach to support a more semantic approach; conversion of metadata created using the current specification would be possible with special treatment of encoding schemes, although it needs to be noted that the special-case heuristics needed for converting encoding schemes would grow as the number of encoding schemes itself grows.
4. For the longer term, the migration to a more semantic model may be considered. This would involve creating vocabularies as separate collections, e.g. defined according to SKOS, and upgrading of the model to reflect more closely the DCMI Abstract Model.

1 Background and scope

The program "Overheid heeft *Antwoord*©" (Government has Answer) has requested a "collegial quality assessment" of the Government Web Metadata Standard (Overheid.nl Web Metadata Standaard – OWMS)

The following questions were specified:

1. To what extent is OWMS semantically in line with Dublin Core semantics?
2. To what extent is OWMS conceptually in line with the DCMI Abstract Model?
3. To what extent is the technical framework suitable to enable a meaningful validation of instance metadata?
4. Which recommendations can be given from a Dublin Core perspective for further development of OWMS?
5. To what extent is the approach of OWMS more widely applicable, for example in a European context and the DC-Government Application Profile?

The review was conducted in the months October and November 2008 by a team of experts (Thomas Baker, Makx Dekkers and Mikael Nilsson) on the basis of a set of documents that describe the semantics, model and technical framework of OWMS:

- http://standaarden.overheid.nl/owms/3.5/doc/. On-line documentation of the properties and encoding of OWMS.
- http://standaarden.overheid.nl/owms/3.5/doc/concepten.pdf. Conceptual model of OWMS.
- http://www.dublincore.org/documents/dcmi-terms/. Semantics of DCMI Metadata Terms.
- http://www.dublincore.org/documents/abstract-model/. DCMI Abstract Model.
- http://standaarden.overheid.nl/owms/3.5/schemas.html.. Overview of technical framework in XML Schema and Schematron.
- http://www.semic.eu/semic/view/index.xhtml. Semantic Interoperability Centre Europe.
- http://www.dublincore.org/dcgapwiki. Description of Dublin Core Application Profile for Government Information
- http://dublincore.org/documents/singapore-framework/. Model for Dublin Core Application Profiles.

2 Approach

As requested in the terms of reference, the review was undertaken in a number of steps that are described in more detail in the following sections.

3 Description of tasks and outcomes

3.1 Task 2: Semantics

3.1.1 Objective

The objective of this task was to verify that the properties defined for the "Core Component" ("kern" and "mantel") were in conformance with the semantics of the DCMI Terms.

3.1.2 Approach

The analysis of the semantics of the terms (properties and encoding schemes) was done by reviewing the descriptions in the on-line documentation of OWMS[1].

The work in this task concentrated on the natural-language semantics of the OWMS terms and identified where definitions were either unclear or where the definitions were not in line with the semantics of the DCMI Terms.

The examples with snippets of HTML and XML were considered insofar as they helped the reviewers better understand the semantics wherever necessary.

3.1.3 Results

Annex A contains the detailed results of the analysis with a number of specific comments related to the properties and encoding schemes,

3.1.4 Conclusion

The conclusion of the work is that the semantics of OWMS 3.5 are by and large in line with the definitions of DCMI Terms. Additional properties are also almost all in line with the approach underlying the DCMI terms. A number of minor issues have been identified in the review that may need further clarification in the documentation.

Two more fundamental issues have been identified.

1. The inclusion of contact details for `dcterms:contributor` violates the one-to-one principle. This could better be handled with separate descriptions of organisations and contact persons, e.g. in a thesaurus.
2. The property `overheidvac:revisionHistory` does not fit the general model of properties and subproperties that Dublin Core is based on, and also is not in line with the OWMS model reviewed under task 3 below. This could be better handled with separate descriptions of revision events with links to the Q&A record.

The reviewers understand that the scope of the review is limited to the "Core Component" of OWMS and that there is no obligation for the extensions to be interoperable outside of the specific application in which they are defined.

[1] http://standaarden.overheid.nl/owms/3.5/doc/

3.2 Task 3: Model

3.2.1 Objective

The objective of this task was to compare the model underlying OWMS with the DCMI Abstract Model.

3.2.2 Approach

To facilitate the comparison, a formal language was developed to represent the constructs of the conceptual model of OWMS on the model of DC-Text[1]. DC-Text defines a syntax for serialising or representing a DC metadata description set as defined in the DCMI Abstract Model in plain text, and OWMS-Text does the same for the OWMS model.

3.2.3 Results

Annex B contains the results of the review of the conformity of the OWMS conceptual model with the Abstract Model.

3.2.4 Conclusion

On the basis of the work done under this task, we conclude that the conceptual model of OWMS is an improved version of the model that is described in the DCMI XML encoding guidelines of 2003[2], rather than being based on the DCMI Abstract Model[3].

One major difference with the DCMI Abstract Model is that OWMS does not attach syntax encoding schemes and languages directly to value strings the way the DCMI Abstract Model does. Moreover, the DCMI Abstract Model does not consider language to be a type of encoding scheme. The model differentiates between syntax encoding schemes and vocabulary encoding schemes in a manner consistent with the DCMI Abstract Model, though the distinction is not reflected in syntax.

The review has looked at the model from the perspective of what would be needed to convert data expressed according to the OWMS model to conform to the DCMI Abstract Model.

For most of the elements in the OWMS model this kind of conversion would be fairly straightforward. In specific cases, however, the correct conversion of data may in practice involve defining and hard-coding specific rules for the transformation of statements using specific properties, and some properties may require special attention or manual intervention.

In the OWMS application profile, for example, `dcterms:subject` can be used with free text (`owms:ValueString`), a pointer (`owms:ValueURI`), a value list (`owms:ValueString` with `owms:VocabularyEncodingScheme`), or a syntax encoding scheme (`owms:ValueString` with `owms:SyntaxEncodingScheme`).

In such cases, the conversion would involve the usage of specific knowledge related to what kind of scheme is being used as explained in Annex B.

[1] http://dublincore.org/documents/dc-text/
[2] http://dublincore.org/documents/dc-xml-guidelines/
[3] http://dublincore.org/documents/abstract-model/

3.3 Task 5: Technical Framework

3.3.1 Objective

The objective of this task was to assess the technical framework[1] with specific attention to its practical applicability.

3.3.2 Approach

The review in this area first looked in detail at the XML framework, reviewing the templates, the structure and the relationship with the conceptual model reviewed in task 3 above. It then looked at the HTML encoding guidelines, reviewing the basic model in comparison with the model underlying the XML framework. As the XML framework is the central aspect of OWMS, the review also looked briefly at future developments in XML.

3.3.3 Results

A detailed description of the results of the review of the Technical Framework is contained in Annex C.

3.3.4 Conclusion

The main conclusion of this task is that the XML framework is adequate for the implementation of OWMS but that it does not completely implement the conceptual model reviewed in task 3. Particularly, the XML encoding does not distinguish between vocabulary and syntax encoding schemes. It may be considered to either modify the technical framework to reflect the conceptual model, or to adapt the conceptual model to reflect the implicit model underlying the XML approach.

The HTML encoding implements the model behind the XML approach but does not allow the identification of the language of the described resource. The reviewers are of the opinion that the documentation of the HTML encoding could be improved.

No immediate actions are recommended regarding new XML developments.

3.4 Task 6: Future developments

3.4.1 Objective

The objective of this task was to assess the relevance of new technical developments in the area of interoperability for OWMS in the medium term.

3.4.2 Approach

The review concentrated on looking at how the OWMS specifications relate to more semantic approaches that are developing, particularly as part of Semantic Web technologies. While XML approaches concentrate on defining shared semantics within a closed group of implementers (which is sometimes referred to as "intra"-operability), Semantic Web approaches (based on RDF as a modeling language) aim at allowing a wider context of open exchange of information ("inter"-operability).

[1] http://standaarden.overheid.nl/owms/3.5/schemas.html

The review considered particularly the mechanisms that are emerging for the use of controlled vocabularies, or vocabulary encoding schemes as they are referred to in the Dublin Core context.

3.4.3 Results

The detailed results of this task are contained in Annex D.

3.4.4 Conclusion

The main conclusion of this task is that while OWMS is essentially a method for generating customized XML languages, the "Core Component" of OWMS should be transformable into RDF with special treatment of encoding schemes.

The controlled vocabularies ("waardelijsten") that are being used in OWMS could be candidates for what is sometimes referred to as "SKOSification" (expression in RDF according to the SKOS[1] specification with assignment of URIs to the individual terms in the vocabulary) which would allow other organizations to use those vocabularies as vocabulary encoding schemes.

Conversely, if OWMS would start referring to controlled vocabulary terms by reference rather than (or in addition to) by including plain-text terms in instance metadata, this would allow OWMS to use external vocabularies such as Eurovoc[2], the European thesaurus of administrative terms, or GEMET [3], the European thesaurus of environmental terminology.

3.5 Task 7: European applicability

3.5.1 Objective

The objective of this task was to examine how OWMS could contribute to international discussions, e.g. in the DCMI Government Community and in platforms for cooperation like the CEN/ISSS eGov-Share Workshop and SEMIC.EU.

3.5.2 Approach

This task was addressed through desk research, direct involvement from reviewers in the activities involved, and one telephone interview. The review looked at the situation in Europe and beyond, identifying some activities that have similarity with the OWMS work, then looked at two specific European initiatives, SEMIC.EU and CEN/ISSS eGov-Share, the global activities in the Dublin Core Metadata Initiative, and considered some international conference series that may have relevance for OWMS.

3.5.3 Results

The results of this task are contained in Annex E.

[1] Simple Knowledge Organization System, http://www.w3.org/2004/02/skos/

[2] http://europa.eu/eurovoc/

[3] http://www.eionet.europa.eu/gemet

3.5.4 Conclusion

The main conclusions of this task are that OWMS could act as a useful example for other countries. For the exchange of experience with others, it would be necessary to provide information about OWMS in English. Liaising with other initiatives and participating in international platforms and conferences could help the OWMS team to learn from experiences elsewhere.

4 Overall conclusions

On the basis of the conclusions of the specific tasks as described above, a number of overall conclusions can be formulated.

1. The semantic definitions and the XML Framework that OWMS is based on are considered to be adequate.
2. There is a problem with the consistency of the modelling, as the conceptual model, the implicit model underlying the XML approach and the model behind the HTML encoding are slightly different. The reviewers recommend that a choice is made for one consistent model and that the application of that model be consistently enforced for the "Core Component" of OWMS.
3. There are no fundamental problems if at some point in the future a decision were to be taken to upgrade the OWMS approach to support a more semantic approach; conversion of metadata created using the current specification would be possible with special treatment of encoding schemes, although it needs to be noted that the special-case heuristics needed for converting encoding schemes would grow as the number of encoding schemes itself grows.
4. For the longer term, the migration to a more semantic model may be considered. This would involve creating vocabularies as separate collections, e.g. defined according to SKOS, and upgrading of the model to reflect more closely the DCMI Abstract Model.

Annex A. Terms used in OWMS 3.5: analysis of natural-language semantics

A.1 Scope of the review

The following review covers the natural-language semantics (labels, definitions, usage comments, and examples) for the terms listed as properties (eigenschappen)[1], syntax encoding schemes (syntax-codeer-schema's)[2], and vocabulary encoding schemes (waardelijsten)[3].

The review also covered several terms that were not listed in the above but were either mentioned in the usage examples or listed on the pages "Toepassing: Decentrale Regelgeving"[4] or "Toepassing: Vraag Antwoord Combinatie"[5].

A.2 General comments

Compared with the DCMI Abstract Model[6], the Conceptual Model OWMS 3.5[7] treats syntax encoding schemes, vocabulary encoding schemes, and language tags as variant types of a single category, "encoding scheme". This part of the review comments on the classification and use of specific encoding schemes in the OWMS specification. The general issue of mapping OWMS encoding schemes to entities of the DCMI Abstract Model (and thus to RDF) is discussed in a separate part of the review (see Annex B, section B.4).

The reviewers note that the terms of vocabulary encoding schemes could be assigned URIs, making them easier to cite with precision in applications (including applications outside the realm of government information). This is also discussed in a separate part of the review (see Annex D, section D.4 and D.5).

Comments on specific terms (see also attached spreadsheet)

1. The entry for `dcterms:accessRights` refers to a term "`http://www.overheid.nl/profile/it/administrator`" for which the reviewers could not find documentation.
2. The properties `dcterms:temporal`, `dcterms:available`, and `dcterms:valid` are used here with the DCMI Period Encoding Scheme, which specifies a DCSV ("structured value") syntax for encoding time intervals in a text string[8]. The OWMS profile, however, recommends that this information be encoded in XML, following a suggestion made in an earlier version of the DCMI Period specification[9]. When this specification was revised in 2006, the reference to alternative XML options was dropped in order to remove ambiguity about the meaning of `dcterms:Period` as a

[1] http://standaarden.overheid.nl/owms/3.5/doc/eigenschappen.html
[2] http://standaarden.overheid.nl/owms/3.5/doc/syntax-codeer-schemas.html
[3] http://standaarden.overheid.nl/owms/3.5/doc/waardelijsten.html
[4] http://standaarden.overheid.nl/owms/3.5/doc/toepassingen/decentrale-regelgeving.html
[5] http://standaarden.overheid.nl/owms/3.5/doc/toepassingen/vraag-antwoord-combinatie.html
[6] http://dublincore.org/documents/abstract-model/
[7] http://standaarden.overheid.nl/owms/3.5/doc/concepten.pdf
[8] http://dublincore.org/documents/dcmi-period/
[9] http://dublincore.org/documents/2000/07/28/dcmi-period/

syntax encoding scheme (i.e., RDF datatype) specifically for the legacy "structured value" syntax. This ambiguity would need to be corrected if OWMS data were converted into RDF triples for the purposes of interoperability – either by converting the XML representation into DCSV syntax or perhaps by representing time intervals as resources.

3. It is unclear what is gained by using `dcterms:coverage` as a "mantel" property in addition to `dcterms:temporal` and `dcterms:spatial`, which are already used in the "kern". The property `dcterms:coverage` does provide a place to put information about jurisdiction, but is that really the intention? As noted in the section "niet te verwarren met", `dcterms:coverage` has a potentially confusing overlap with `dcterms:subject`.
4. In the "`overheidvac:`" namespace, three properties (`prioriteit`, `statusRedactie`, and `topN`) are defined in parallel to three vocabulary encoding schemes with the same names in uppercase (`Prioriteit`, `StatusRedactie`, and `TopN`). URIs are case-sensitive, so this practice is valid in a technical sense. It is worth noting, however, that DCMI has decided to avoid coining URIs that differ only with respect to case because of the potential for confusion. Arguments can be made either way on this issue.
5. The property `dcterms:educationLevel` (Audience Education Level) is defined as "A class of entity, defined in terms of progression through an educational or training context". The intention is to characterize education level in terms of the people who represent that level, such as "sixth-grade students". It is worth noting that the Dutch definition for "`Opleidingsniveau`" ("Omschrijving van het opleidingsniveau van de doelgroep") does not reflect this subtle distinction. It was indeed precisely to remove a long-recognized ambiguity in this regard that the current definition was published in January 2007. The reviewers do not believe this impacts the values provided for the property (such as "ISCED 0"), though it does affect how those values are interpreted. It might be worth double-checking this point with colleagues in the DCMI Education Community.

Two more fundamental issues have been identified.

1. The inclusion of contact details for `dcterms:contributor` violates the one-to-one principle. According to this principle, the value of the metadata for contributor identifies the person or organization that contributed, while a description of the person or organizations should be contained in a separate metadata record, much like the use of "name authority files" in library and information science. Apart from the fundamental issue that this information is not *about* the primary resource being described, a practical issue of including this kind of information in instance metadata is that when contact details change (moving of office, changes in e-mail addresses) all metadata records that contain these details would need to be updated.
2. The property `overheidvac:revisionHistory` does not fit the general model of properties and subproperties that Dublin Core is based on, and also is not in line with the OWMS model reviewed under task 3 below. Subproperties are supposed to "refine" a property in such a way that the value of a subproperty is also a valid value of the property. In the case of `revisionHistory`, the "subelements" `version`, `dateTime`, `author` and `comment` are not proper refinements of the property. In effect, this information appears to be a description of another "resource", namely an event in a revision history. These events could be separately described with a link to the Q&A record that was revised.

A.3 Conclusion

The semantics of OWMS 3.5 are by and large in line with the definitions of DCMI Terms. Additional properties are also almost all in line with the approach underlying the DCMI terms. A number of minor issues have been identified in the review that may need clarification in the documentation. The two major issues identified above need further consideration.

OWMS review

Task : Analysis of semantics

No.	Namespace	Property	Comments
	Core		
1	dcterms	identifier	
2	dcterms	title	
3	dcterms	type	
4	dcterms	language	
5	dcterms	creator	
6	dcterms	modified	The difference between dcterms:modified and dcterms:created is a bit difficult to understand.
7	dcterms	temporal	The difference between dcterms:temporal and dcterms:valid may not be easy to explain. See also "Comments on specific terms", point 2.
8	dcterms	spatial	The reference to overheid:PostcodeHuisnummer is incorrect. It is a Syntax Encoding Scheme, and the correct link is http://standaarden.overheid.nl/owms/3.5/doc/syntax-codeer-schemas/overheid.postcodehuisnummer.html.
	Mantle		
9	dcterms	abstract	
10	dcterms	accessRights	
11	dcterms	alternative	The comment suggests that a standard mechanism could be devised to make abbreviations somehow consistent across collections.
12	dcterms	audience	
13	dcterms	available	See "Comments on specific terms", point 2.
14	dcterms	conformsTo	
15	dcterms	contributor	The usage guidelines say that contact details need to be included if not available elsewhere in the metadata. Aside from violating the one-to-one principle (according to which contact details belong in a separate description), this could muddy the metadata with details that cannot properly be parsed.
16	dcterms	coverage	See "Comments on specific terms", point 3.
17	dcterms	created	
18	dcterms	date	
19	dcterms	dateAccepted	There is no further comment to explain how this property might be used.
20	dcterms	dateCopyrighted	
21	dcterms	dateSubmitted	
22	dcterms	description	
23	dcterms	educationLevel	See "Comments on specific terms", point 5.
24	dcterms	extent	
25	dcterms	format	
26	dcterms	hasFormat	
27	dcterms	hasPart	
28	dcterms	hasVersion	
29	dcterms	instructionalMethod	
30	dcterms	isFormatOf	
31	dcterms	isPartOf	
32	dcterms	isReferencedBy	
33	dcterms	isReplacedBy	
34	dcterms	isRequiredBy	
35	dcterms	isVersionOf	
36	dcterms	issued	
37	dcterms	license	
38	dcterms	mediator	
39	dcterms	provenance	
40	dcterms	publisher	
41	dcterms	references	
42	dcterms	relation	
43	dcterms	replaces	
44	dcterms	requires	
45	dcterms	rights	
46	dcterms	rightsHolder	
47	dcterms	source	
48	dcterms	subject	
49	dcterms	tableOfContents	
50	dcterms	valid	The difference between dcterms:temporal and dcterms:valid may not be easy to explain. See also "Comments on specific terms", point 2.
51	overheid	abbreviation	The usage comment for dcterms:alternative also refers to abbreviations, so it is not clear why this additional property is needed. It is also unclear whether this intended to be a sub-property of dcterms:title.
52	overheid	authority	
53	overheid	isRatifiedBy	The reviewer assumes this is a sub-property of dcterms:relation?
	Vocabulary Encoding Schemes		
54	overheid	Doelgroep	
55	overheid	Informatietype	
56	overheid	TaxonomieBeleidsagenda	
57	overheid	Adviescollege	
58	overheid	Deelgemeente	
59	overheid	DienstAgentschapInstellingOfProject	
60	overheid	Gemeente	
61	overheid	HoogCollegeVanStaat	
62	overheid	Koepelorganisatie	
63	overheid	KoninklijkHuis	
64	overheid	Ministerie	
65	overheid	OpenbaarLichaamVoorBedrijfEnBeroep	
66	overheid	Politiekorps	
67	overheid	Provincie	
68	overheid	RechterlijkeMacht	
69	overheid	Regering	
70	overheid	RegionaalSamenwerkingsorgaan	
71	overheid	StatenGeneraal	
72	overheid	Waterschap	
73	overheid	ZelfstandigBestuursorgaan	
74	overheid	BestuursorgaanDeelgemeente	
75	overheid	BestuursorgaanGemeente	
76	overheid	BestuursorgaanMinisterie	
77	overheid	BestuursorgaanProvincie	
78	overheid	BestuursorgaanRegionaalSamenwerkingsorgaan	
79	overheid	BestuursorgaanWaterschap	
	Syntax Encoding Schemes		
80		VrijeTekst	
81	dcterms	IMT	
82	dcterms	Period	
83	dcterms	RFC4646	
84	dcterms	URI	
85	dcterms	W3CDTF	
86	overheid	PostcodeHuisnummer	Listed (correctly) as SES, but cited (incorrectly) in the entry for dcterms:spatial as VES.
	Terms not listed in the above lists		
87	overheidvq	DocumentType	This is mentioned in the example for dcterms:type.
88	overheidvac	datumcontrole	See http://standaarden.overheid.nl/owms/3.5/doc/eigenschappen/overheidvac.datumcontrole.html. This property is not defined.
89	overheidvac	revisionhistory	See http://standaarden.overheid.nl/owms/3.5/doc/eigenschappen/overheidvac.revisionhistory.html. This would appear to be a pointer to a separate description of a revision history, but the reference to "sub-elements" is unclear.
90	overheidvac	contentaanbieder	Ok.
91	overheidrg	BestuursorgaanEilandgebied	Listed in "Decentrale Regelgeving".
92	overheidrg	BestuursorgaanNederlandseAntillen	Listed in "Decentrale Regelgeving".
93	overheidrg	Eilandgebied	Listed in "Decentrale Regelgeving".
94	overheidrg	NederlandseAntillen	Listed in "Decentrale Regelgeving".
95	overheidvac	BedrijfContentAanbieder	Listed in "Vraag Antwoord Combinatie". Has no controlled terms!
96	overheidvac	Prioriteit	Listed in "Vraag Antwoord Combinatie". URI is identical to overheidvac:prioriteit (#101) except for case!
97	overheidvac	StatusRedactie	Listed in "Vraag Antwoord Combinatie". Fine as is, though could also be modeled as Syntax Encoding Scheme. URI is identical to overheidvac:statusRedactie (#103) except for case!
98	overheidvac	TopN	Listed in "Vraag Antwoord Combinatie". Fine as is, though could also be modeled as Syntax Encoding Scheme. URI is identical to overheidvac:topN (#104) except for case!
99	overheidvac	contentaanbieder	Listed in "Vraag Antwoord Combinatie".
100	overheidvac	datumControle	Listed in "Vraag Antwoord Combinatie".
101	overheidvac	prioriteit	Listed in "Vraag Antwoord Combinatie". URI is identical to overheidvac:Prioriteit (#96) except for case! See also "Comments on specific terms", point 4
102	overheidvac	revisionHistory	Listed in "Vraag Antwoord Combinatie".
103	overheidvac	statusRedactie	Listed in "Vraag Antwoord Combinatie". URI is identical to overheidvac:StatusRedactie (#97) except for case! See also "Comments on specific terms", point 4.
104	overheidvac	topN	Listed in "Vraag Antwoord Combinatie". URI is identical to overheidvac:TopN (#98) except for case! See also "Comments on specific terms", point 4.

Annex B. Conformity of Conceptual Model OWMS 3.5 with the DCMI Abstract Model

B.1 Scope of review

This part of the review compares the Conceptual Model OWMS 3.5, page 8] with the Description Set Model, a part of the DCMI Abstract Model. The review also compares the OWMS model to a similar description model specified in "Guidelines for encoding Dublin Core metadata in XML" of 2003[1], as the relationship between this earlier model and the DCMI Abstract Model has been well-described in a DCMI note[2].

B.2 Method of review

The draft document "Interoperability levels for Dublin Core metadata"[3] distinguishes four levels of interoperability. The most basic level (Level 1, "Shared Terms Defined") involves the use of shared concepts as defined in natural language. The conformity of the OWMS specification on this level is discussed in another part of this review (see Annex A). Using criteria defined in the interoperability levels document, this comparison of the OWMS model with the Description Set Model of DCMI Abstract Model asks the following questions:

- Is there a complete mapping between elements of the OWMS model to elements used in the DC-TEXT format[4]? DC-TEXT is a simple text format for representing a Dublin Core metadata Description Set. In the interoperability level model, this conformance test corresponds to Level 3, "Description Set Syntactic Interoperability".
- Can data based on the OWMS model can be completely mapped to RDF triples in a way that respects the declared domains and ranges of the properties used? In the interoperability level model, this conformance test corresponds to Level 2, "Formal Semantic Interoperability".

In order to present the mappings clearly and unambiguously, the examples given below use the DC-TEXT format alongside an improvised "OWMS-TEXT format". The OWMS-TEXT format was created by using labels of elements in the UML diagram on page 8 of the Conceptual Model specification, in camel-case, as names for OWMS-TEXT elements. To avoid confusion, these elements (and the corresponding DC-TEXT elements) are cited in the text using qualified names (e.g., `owms:Description` for the OWMS-model element Description and `dctext:Description` for the DCMI Abstract Model element Description).

It was easy to invent this format because the OWMS model uses the same basic hierarchy as the DCMI Abstract Model. Accordingly, OWMS-TEXT follows the same basic nesting conventions as DC-TEXT with one important exception: in OWMS-TEXT, the elements `owms:Language` and `owms:SyntaxEncodingScheme` are nested within an enclosing element `owms:Statement`, whereas in DC-TEXT the analogous elements `dctext:Language` and `dctext:SyntaxEncodingSchemeURI` are nested within either a

[1] http://dublincore.org/documents/dc-xml-guidelines/
[2] http://dublincore.org/documents/dc-ds-xml-notes/
[3] http://dublincore.org/documents/interoperability-levels/
[4] http://dublincore.org/documents/dc-text/

`dctext:ValueString` or `dctext:LiteralValueString` element. The complete list of OWMS-TEXT elements is:

```
owms:DescribedResource
owms:Description
owms:Statement
owms:Property
owms:ValueURI
owms:ValueString
owms:EncodingScheme
owms:VocabularyEncodingScheme
owms:SyntaxEncodingScheme
owms:Language
```

Note that no element is needed for the OWMS-model element "value surrogate", just as in DC-TEXT there are no corresponding elements for literal and non-literal value surrogates. As DC-TEXT and OWMS-TEXT are both essentially shorthand conventions which focus on mappable syntactic elements and omit intermediate modeling elements – notably the value surrogates – their use helps keep this review concise and to-the-point.

B.3 Comparison of the OWMS model with the 2003 DC-XML model

The OWMS model is similar to the description model defined in "Guidelines for encoding Dublin Core metadata in XML" of 2003, a predecessor of the DCMI Abstract Model of 2007. It is useful to start by comparing the OWMS model to this older "abstract model" because the differences between the 2003 and 2008 models, and their consequences for mapping, are discussed in some detail in a DCMI note referred to in section B.1 above.

The 2003 DC-XML specification outlines the following description model:

- A qualified DC record is made up of one or more properties and their associated values.
- Each property is an attribute of the resource being described.
- Each property must be either:
  - one of the 15 DC elements,
  - one of the other elements recommended by DCMI (e.g. audience)
  - one of the element refinements listed in the DCMI Metadata Terms recommendation
- Properties may be repeated.
- Each value is a literal string.
- Each value may have an associated encoding scheme.
- Each encoding scheme has a name.
- Each literal string value may have an associated language (e.g. en-GB).

The OWMS model differs from this 2003 model in the following ways:

- The OWMS model allows the use of properties from non-DCMI namespaces which follow the same model (e.g., `overheid:`)
- The OWMS model distinguishes value URIs from literal strings.
- The OWMS model distinguishes between syntax encoding schemes and vocabulary encoding schemes.

- In the OWMS model, the language associated with literal string values is considered to be a type of encoding scheme.

With the exception of the final point, these differences improve on the 2003 model, bringing the OWMS model closer to the Description Set Model defined the DCMI Abstract Model.

The OWMS model and the DCMI Abstract Model share the following analogous constructs:

owms:DescribedResource	dctext:ResourceURI
owms:Description	dctext:Description
owms:Statement	dctext:Statement
owms:Property	dctext:PropertyURI
owms:ValueURI	dctext:ValueURI

The OWMS model differs from the DCMI Abstract Model as follows:

- The OWMS model has no construct analogous to dctext:DescriptionSet, which in practice means that metadata records are limited to describing a single resource.
- The OWMS model has no concept analogous to the notion that a non-literal value surrogate may include multiple value strings, with the practical consequence that value strings in multiple languages cannot be grouped together in a single statement.
- The OWMS model has a single construct owms:ValueString, whereas the DCMI Abstract Model distinguishes dctext:LiteralValueString (for a Value String in a Literal Value Surrogate) and dctext:ValueString (for a Value String in a Non-Literal Value Surrogate):

owms:ValueString	dctext:ValueString or dctext:LiteralValueString

  It is worth noting that the DCMI Abstract Model distinguishes between literal and non-literal values because these are modeled and expressed using different and incompatible patterns of RDF triples, whereby non-literal values (unlike literal values) may themselves be the subject of further description.

- Analogously to the DCMI Abstract Model, the OWMS model distinguishes between vocabulary encoding schemes, syntax encoding schemes, and languages:

owms:VocabularyEncodingScheme	dctext:VocabularyEncodingSchemeURI
owms:SyntaxEncodingScheme	dctext:SyntaxEncodingSchemeURI
owms:Language	dctext:Language

  Unlike the DCMI Abstract Model, however, OWMS does not attach syntax encoding schemes and languages directly to value strings – a fact that is reflected in how the OWMS-TEXT examples below are nested. (Note also that the DCMI Abstract Model does not consider language to be a type of encoding scheme – a point which is however of no practical consequence for mapping.)

  The discussion of mapping below assumes that these three types of encoding scheme are properly distinguished in the concrete implementation syntaxes. Whether this is the case for the syntaxes defined for the OWMS package is considered in another part of this review (see Annex C).

B.4 Mappings between the OWMS model and DCMI Abstract Model

In order to respect the declared range semantics of RDF properties, mappings need to distinguish between properties with a non-literal range and properties with a literal range.

OWMS-conformant data	*DCAM-conformant data (DC-TEXT)*
owms:ValueString owms:Statement (owms:Property (dcterms:date) owms:ValueString ("2008-10-18"))	**dctext:LiteralValueString** Statement (PropertyURI (dcterms:date) LiteralValueString ("2008-10-18"))
owms:ValueString owms:Statement (owms:Property (dcterms:subject) owms:ValueString ("fraude"))	**dctext:ValueString** Statement (PropertyURI (dcterms:subject) ValueString ("fraude"))
owms:VocabularyEncodingScheme owms:Statement (owms:Property (dcterms:subject) owms:VocabularyEncodingScheme (overheid:Thesaurus) owms:ValueString ("fraude"))	**dctext:VocabularyEncodingSchemeURI** Statement (PropertyURI (dcterms:subject) VocabularyEncodingSchemeURI (overheid:Thesaurus) ValueString ("fraude"))
owms:SyntaxEncodingScheme owms:Statement (owms:Property (dcterms:language) owms:ValueString ("en") owms:SyntaxEncodingScheme (dcterms:ISO639-2)) owms:Statement (owms:Property (dcterms:date) owms:ValueString ("1976") owms:SyntaxEncodingScheme (dcterms:W3CDTF))	**dctext:SyntaxEncodingSchemeURI** Statement (PropertyURI (dcterms:language) ValueString ("en" SyntaxEncodingSchemeURI (dcterms:ISO639-2))) Statement (PropertyURI (dcterms:date) LiteralValueString ("1976" SyntaxEncodingSchemeURI (dcterms:W3CDTF)))
owms:Language owms:Statement (owms:Property (dcterms:title) owms:ValueString ("My Life") owms:Language (en-GB)) owms:Statement (owms:Property (dcterms:subject) owms:ValueString ("fraud") owms:Language (en-GB))	**dctext:Language** Statement (PropertyURI (dcterms:title) LiteralValueString ("My Life" Language (en-GB))) Statement (PropertyURI (dcterms:subject) ValueString ("fraud" Language (en-GB)))

Note the dependency on range, in particular, for the proper mapping of `owms:ValueString`, `owns:SyntaxEncodingScheme`, and `owms:Language`. The correct conversion of data may in practice involve defining and hard-coding specific rules for the transformation of statements using specific properties, and some properties may require special attention or manual intervention.

In the OWMS application profile, for example, `dcterms:subject` can be used with free text (`owms:ValueString`), a pointer (`owms:ValueURI`), a value list (`owms:ValueString` with `owms:VocabularyEncodingScheme`), or a syntax encoding scheme (`owms:ValueString` with `owms:SyntaxEncodingScheme`).

B.5 Conclusion

In conclusion, it needs to be noted that the current model is not in line with the DCMI Abstract Model but that instance metadata that are based on the OWMS model, as currently defined, can be converted to DCMI Abstract Model-conformant instance metadata with the proviso that some special processing may be necessary.

The reviewers would like to add that care needs to be taken when the OWMS model is further developed as there is a risk that extensions of the OWMS model could create a situation where such a mapping is impossible.

It may be beneficial on the longer term to consider modifying the OWMS model to bring it more in line with the DCMI Abstract Model.

Annex C. Technical framework

C.1 Overview of analysis

This part of the review covers the technical framework, comprising an XML syntax, the corresponding XML schemas, and an HTML syntax.

The OWMS technical framework provides two syntaxes, one using a custom XML template-based language and one using HTML `<meta>` elements. Both are loosely based on the corresponding DC expressions, but due to fundamentally different internal structure, they need to be reviewed independently.

The framework has been reviewed for internal consistency, compatibility with XML and HTML software, coherence with the conceptual model, and compatibility with Dublin Core.

C.2 XML

C.2.1 Templating

The XML specification is, in effect, a template for designing application-specific XML languages, designed to support local customization through combination of generic and application specific XML Schema fragments. Through a system of XML Schema inclusions, custom languages can be developed, while retaining the basic structure of the resulting XML documents. Interoperability is achieved through the use of a common metadata "core" which we refer to as the "Core Component" that consists of a "kern" and a "mantel" of metadata elements or properties.

The review has not uncovered any issues on the XML level with the template approach, and has found the documentation to be clear and adequate.

We have also seen no reason to criticise the XML templating functionality, and our analysis shows that the templating approach is, in itself, not a source of interoperability issues. It is difficult to assess whether the templating approach is suitable for deployment without knowing more details about the intended use of the framework, but it presents no inherent issues.

C.2.2 Structure

The XML syntax uses a relatively unorthodox scaffolding structure for metadata syntaxes, separating common "kern" and "mantel" elements from application-specific elements. The XML languages have a common structure, which is essentially

```
<[app-root]>
   <meta>

      <owmskern>
      </owmskern>

      <owmsmantel>
      </owmsmantel>
```

```
      <[app]meta>
      </[app]meta>

      <othermeta>
      </othermeta>

   </meta>

   <body>
   </body>

</[app-root]>
```

[app-root] is an application-specific root XML element, while [app]meta is an application specific metadata container element. [app-root] and [app] are supposed to be replaced by application-specific names.

Within each of the four metadata containers, metadata elements may appear as follows:

- The owmskern contains mandatory elements, and may not be modified.
- The owmsmantel contains a set of preselected but optional elements that may be included or excluded from an application.
- [app]meta contains arbitrary local application-specific elements
- othermeta contains arbitrary imported elements from other namespaces.

It is the stated intention of the OWMS designers that only elements within the owmskern and owmsmantel containers (the "Core Component") follow the OWMS conceptual model. Elements in the [app]meta and othermeta containers are out of scope for conformance to the conceptual model. In consequence, only the "Core Component" needs review for compatibility with the conceptual model.

The review did not find any issues with the multi-container structure or the XML Schemas used to implement this structure. In fact, given the stated goal of controlling the contents of the "Core Component" while leaving the [app]meta and othermeta containers uncontrolled, the given structure is fully adequate. The separation makes is easier to construct XML schemas, and easier to design processors that handle just the "Core Component", while retaining the possibility of extending the model. The XML Schemas adequately implement validation of this overarching structure.

C.2.3 Relationship with the conceptual model

Conformance to the conceptual model is maintained on an element-by-element basis within the "Core Component". The [app]meta and othermeta containers are, as previously stated, essentially uncontrolled.

The interpretation of an XML document following the OWMS framework is as a single Description, with a set of Statements conforming to the OWMS conceptual model, one statement for each element within the owmskern and owmsmantel containers. The element in the [app]meta and othermeta containers cannot without further knowledge be interpreted as Statements.

C.2.3.1 Element structure within the "Core Component"

All "Core Component" elements use a single string value (XML content), with the exception of `dcterms:temporal/available/valid`, which all use an XML sub-structure containing `<start>` and `<end>` elements, as per earlier versions of the DCMI Period Encoding Scheme.

There are three attributes used, as appropriate, for the elements:

- `"xml:lang"`, for encoding languages
- `"resourceIdentifier"`, for giving URIs for referenced information resources
- `"scheme"`, for giving a vocabulary or syntax encoding scheme

Each element is restricted to an `xsd complexType` or `simpleType`. Each such type restricts the syntax of the string value and gives the allowed attributes.

For example, the element `owms:isRatifiedBy` is defined with a simple string value as content, and only the attribute `"scheme"` is allowed. This attribute is restricted to an enumerated list of vocabulary encoding schemes. However, the fact that the sting value in the metadata is indeed a valid term from the vocabulary used cannot be validated; this is a known limitation of XML Schema.

C.2.3.2 Vocabulary encoding scheme customization

The XML Schema templating mechanism allows for the redefinition of allowed vocabulary encoding schemes for a given element. New schemes are introduced by referencing new vocabulary encoding schemes in overloaded `xsd` types in the application-specific XML schemas.

C.2.3.3 Evaluation

The XML structure of the "Core Component" elements follows the OWMS conceptual model, with two exceptions:

- There is no syntactic separation between syntax and vocabulary encoding schemes. This has consequences for the interpretation (especially the unambiguity) of OWMS metadata; this issue has been covered in a separate part of the review (see Annex B, section 4).
- `dcterms:temporal/available/valid` allow for an XML substructure which is not covered by the conceptual model. While this also calls for element-specific processing, there is no resulting unambiguity.

One minor issue is the overly restrictive use of `"resourceIdentifier"`, i.e. `valueURIs`. From the XML schemas, it is clear that `valueURIs` are only used when the value is a web document or web site. There is no possibility of using URIs for vocabulary terms, to identify individuals or other uses foreseen by the definition of `valueURIs` in the DCMI Abstract Model.

In our opinion, this restriction does not add value but instead excludes possibilities of reusing external URI-identified vocabularies and other URIs.

C.3 HTML

C.3.1 Model

The syntax uses a simple model of <property, scheme, value>, encoded using the <meta> HTML element, with support for namespaces declared in a <link> element.

C.3.2 Comparison with the XML model

- The HTML model shares the issue of not separating syntax and vocabulary encoding schemes.
- There is no distinction between "Core Component" and application-specific extensions. It is not clear from the documentation if only "Core Component" metadata should be used in HTML.
- There is no language support. There is an appropriate HTML "lang" attribute that could be used.
- There is no valueURI support. In comparison, DC-HTML uses a different HTML element for statements with value URIs (and, indeed, for all non-literal properties), i.e. the <link> element.

C.3.3 Interoperability issues

It is unclear whether DC-HTML compatibility is a goal of the HTML syntax. However, the model is not compatible with the DC-HTML model, for the following reasons:

- The profile attribute, required by DC-HTML, is not used.
- DC-HTML does not support Vocabulary encoding schemes, whereas the OWMS encoding allows the specification of both syntax and vocabulary encoding schemes in the scheme attribute (as the OWMS model does not distinguish between the two types of encoding schemes).
- DC-HTML encodes non-literals using the <link> element, whereas the OWMS encoding only uses <meta> elements.

C.3.4 Other issues

The syntax examples uses "." as separator, as in OVERHEID.Informatietype, but the text uses ":".

C.3.5 Conclusion

The HTML syntax is very briefly described, and definitely lacking in detail. If an HTML syntax is necessary, the reviewers recommend to either:

- Harmonize the model more closely with the XML model and the conceptual model, or
- Adopt the DC-HTML model.

In both cases, the model and its relationship to the XML model needs to be more thoroughly documented to avoid confusion.

C.4 Future developments of XML

The reviewers have been asked to assess whether more recent XML developments should influence the design of the technical framework. The reviewers believe the short answer to be no.

There are other ways of performing XML validation (instead of using XML Schema), such as RelaxNG and Schematron. Both of these are capable validation languages, both generally accepted as more capable than XML Schema. However, for the needs of the OWMS framework, XML Schema is clearly sufficient for the current needs, while at the same time enjoying a broad range of tool support.

This should not be taken as a recommendation to avoid either of these alternative schema languages, but instead as a recommendation to let any such considerations be driven by clearly identified needs and deficiencies in the current XML Schemas.

It is also important to note that schema languages can be replaced without affecting the resulting XML languages, allowing for experimentation and even coexistence of several schema languages for the same framework. The reviewers for this reason strongly recommend continuing to keep a clear separation between the design of the XML structure, and the implementation of that structure in a schema language.

XML (and HTML) is also undergoing future developments. Given the great variety and stability of the current set of mark-up tools, the reviewers see no need to consider new versions of XML or HTML for at least another five to ten years.

C.5 Overall Conclusions

The major conclusions from Task 5 are:

- The XML framework is for the most part perfectly functional, but suffers from a few minor incoherencies with the conceptual model.
- The OWMS "kern" and "mantel" can be interpreted in a Dublin Core interoperable way, with the caveats mentioned in the review of the conceptual framework.
- Extensions are not interoperable
- The HTML syntax suffers from incomplete documentation, and needs to be reworked.
- No immediate actions are recommended regarding new XML developments

Annex D. Future developments

D.1 Current situation

In order to maximize uptake, OWMS has a requirement for the simplest possible XML data, and the OWMS metadata model reflects the simplicity of the desired data structures.

As we understand it, the OWMS approach is essentially a method for generating customized XML languages with, one might say, a "Core Component" and "Extension Slots".

The "Core Component" uses mandatory "kern" elements of Dublin Core with optional "mantel" elements from DCMI Metadata Terms.

The "Core Component" follows the Conceptual Model OWMS 3.5, which is similar to the model used by the Dublin Core Metadata Initiative prior to the DCMI Abstract Model of 2005 inasmuch it is limited to the description of one resource and does not distinguish syntactically between Syntax Encoding Schemes (RDF datatypes) and Vocabulary Encoding Schemes. Exceptions to the model are noted, such as the use of XML substructure in values for `dcterms:temporal`, `dcterms:available`, and `dcterms:valid`.

The Extension Slots, then, can be used for any additional XML content that is locally important for a given implementation without regard for the Conceptual Model. In the absence of a coherent model one would expect metadata in the Extension Slots to be too unmanageably diverse to be interoperable outside the boundaries of particular implementations. Enforcing the compatibility of extension metadata may not, however, be practical or even necessary.

If this picture is accurate, then we assume that the efforts of the central OWMS team can focus on enforcing the coherence of "Core Component" metadata. The "intra"-operability of core metadata within the OWMS community is ensured by the Conceptual Model, so this review examines the "inter"-operability of OWMS data with non-OWMS data in future merged-data applications.

D.2 Exposing OWMS data in RDF

At the current state of development, the interoperability of metadata outside of a known application environment depends, in practical terms, on how the data can be expressed as RDF triples. Data expressed as RDF triples is merged simply by aggregating the data sources.

OWMS "Core Component" metadata in XML or HTML should be transformable into RDF triples – this is typically done using the XML stylesheet language XSLT – as long as certain constructs are given special treatment. The special case of encoding schemes is discussed below; other special cases are noted in the analyses for Task 2 and Task 3.

If robust heuristics for transforming the special cases can be defined, then the process of extracting RDF triples from "Core Component" metadata can be automated and made available on demand using the specification Gleaning Resource Descriptions from

Dialects of Languages (GRDDL[1]). GRDDL is a method for associating an XML or HTML document with a transformation recipe, such that a user can extract triples from the structured document with a mouse click.

If future versions of the OWMS model were to clarify the distinction between Vocabulary and Syntax Encoding Schemes, RDF expressions might be supported in addition to the XML syntax.

For example, the recently announced W3C Recommendation RDFa[2] (for "RDF-in-attributes") allows RDF metadata to be embedded in HTML pages using special tag attributes that are normally hidden from display. Using browser plug-ins, however, users would be able to view the underlying metadata or cut-and-paste metadata into their own applications, while also scripts could be developed to automatically extract metadata from documents.

D.3 Transforming encoding schemes

The thorniest problem for automated transformation of OWMS metadata into triples is that of encoding schemes. While the OWMS Conceptual Model seems to follow the DCMI Abstract Model in distinguishing Vocabulary Encoding Schemes from Syntax Encoding Schemes, this distinction is not supported in the intentionally simple XML syntax. Syntactically indistinguishable in the source data, these two types of encoding scheme are expressed differently in RDF triples – Syntax Encoding Schemes as RDF datatypes associated with literals and Vocabulary Encoding Schemes as resources related to values using a `dcam:memberOf` relationship. These mappings are further discussed in the analysis for Task 3.

While the number of encoding schemes supported for OWMS core metadata seems manageably small – currently seven syntax encoding schemes and 27 value lists – the trade-off between simplicity of encoding and complexity of transforms would shift if that list were to grow over time (for example, by using European encoding schemes). Adding encoding schemes in the future would at any rate entail not just a revision of the specification but also the addition of special-case handling to the transformation algorithm. If the case could be made politically, future work might be simplified if the XML encoding for OWMS metadata were revised today to support this distinction syntactically.

D.4 Coining URIs for value vocabulary terms

Currently, both types of encoding scheme in the OWMS Conceptual Model are expressed as value strings. However, the real value of controlled vocabularies for linking data is unleashed when the individual terms of those vocabularies are themselves citable with URIs. The twenty-seven value lists used in OWMS core metadata, for the most part lists of geographic entities and jurisdictions, are good candidates for "SKOSification" – definition as concept schemes described in RDF with concepts identified using URIs.

For example, the thousands of headings of the Library of Congress Subject Headings (LCSH) have been made citable with URIs. Metadata citing those URIs can efficiently and precisely be linked, creating clusters of resources about, for example, Chinese

[1] GRDDL specification http://www.w3.org/2004/01/rdxh/spec

[2] RDFa Syntax http://www.w3.org/2006/07/SWD/RDFa/syntax/

history or Masonic temples. In the case of OWMS vocabularies, the URIs for "Provincie Groningen" or "Ministerie van Justitie" would be available for users outside of the OWMS implementation community. The example of LCSH is good to emulate also because it implements good-practice guidelines for publishing RDF vocabularies on the Web using content negotiation, such that Web users can retrieve readable Web pages while RDF applications can retrieve RDF schemas.

D.5 Using external vocabularies

From another perspective, citing vocabulary terms with URIs is expected to be also possible for European vocabularies such as GEMET and Eurovoc. GEMET, the European standard vocabulary for environmental terminology, is already available in SKOS form[1]. In 2007, a study at ITTIG-CNR proposed to apply SKOS as a mapping methodology to Eurovoc[2], the European Commissions multilingual thesaurus of terms that are relevant for the European institutions.

When such European, multilingual vocabularies allow their terms to be cited by URI, OWMS could point to the concepts in those vocabularies, rather than copy the text of the terms in instance metadata. An immediate advantage of this is that, because the concepts in those vocabularies are expressed in multiple languages, the instance metadata would point to the set of synonyms that are related to the concepts.

[1] GEMET Web services http://www.eionet.europa.eu/gemet/webservices?langcode=en

[2] S. Faro, E. Francesconi, V. Sandrucci. D1.5 – Thesauri KOS analysis and selected thesaurus mapping methodology on the project case-study. Final Version, October 16th, 2007. TENDER No. 10118 - EUROVOC Studies LOT2. ITTIG-CNR – Institute of Legal Information Theory and Techniques. http://dossierdoc.typepad.com/descripteurs/files/presentationd1.5%20Mapping%20Methodology.pdf

Annex E. European applicability

E.1 OWMS in a European context

OWMS as an activity to build an interoperable framework for the Dutch public sector is not unique in Europe or, indeed, the world. In many countries, similar activities can be identified. Examples are:

- Germany: XÖV Framework[1] and SAGA 4.0[2]
- United Kingdom: eGovernment Metadata Standard[3]
- Canada: Government of Canada Metadata Framework[4]
- United States: Webcontent.gov metadata guidelines[5]

Many of those national initiatives are looking for experiences and tools to share with other similar initiatives in other countries.

The IDABC[6] (*Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens*) programme of the European Commission intends to play a coordinating role in the area of Europe-wide interoperability.

OWMS could play a role as a case study in the development of interoperability across Europe. Of course, one of the main issues to assess is whether an approach based on OWMS could be used in other countries. In effect, there does not seem to be a fundamental reason for OWMS not to be applicable in other countries, with another set of controlled vocabularies as appropriate.

However, the fact that OWMS documentation is mainly in Dutch does currently constitute a barrier for wider distribution. If there is a willingness to share experiences with others, translation of the main documentation should be considered.

E.2 European initiatives

E.2.1 SEMIC.EU

A major initiative in Europe at the moment is SEMIC[7] (*Semantic Interoperability Centre Europe*), an initiative funded by the European Commission under the IDABC programme.

The objective of SEMIC.EU is to provide a service to the community of eGovernment programmes in Europe where so-called interoperability assets can be published for sharing and re-use among members states. These assets can include all kinds of documents and tools that help create semantic interoperability across Europe, for example XML schemas, vocabularies and ontologies. These assets are collected, quality

[1] http://tinyurl.com/6lo3tk
[2] http://gsb.download.bva.bund.de/KBSt/SAGA/SAGA_v4.0.pdf
[3] http://www.govtalk.gov.uk/schemasstandards/metadata.asp
[4] http://www.lac-bac.gc.ca/metaforum/014005-03213-e.html
[5] http://www.usa.gov/webcontent/managing_content/organizing/metadata.shtml
[6] http://ec.europa.eu/idabc
[7] http://www.semic.eu/semic/

checked and made available for download in the Web repository. The SEMIC Web site is designed to be a reference source for semantic interoperability in Europe.

In a conference call and subsequent correspondence with Sebastian Sklarß of]init[AG, the company that runs SEMIC.EU, we received the following list of what they may offer:

1. Clearing and coaching towards eGovernment projects (making expertise out of uploaded assets, create awareness for recommendations resulting on pan-European usage)
2. Providing infrastructure (the repository with versioning, release control and "rights" management will be heavily improved in the next months and years)
3. Enabling networking and collaboration (present projects on Web site and in the forum), yearly conference in Brussels[1], community workshops in specific domains. At the moment there are Communities for "eJustice", "Environment", "Methodology", "Taxonomy" and a "EURES" (European Employment Services) Interest Group.

The OWMS properties, vocabularies, model and XML schemas could all be submitted to SEMIC as "Interoperability Assets". The main requirement for registration is that documentation is provided in English. A project fact sheet for submission to SEMIC can be obtained at contact@semic.eu.

SEMIC has published an overview of their assessment process, called "The Vision of the Clearing Process"[2] that describes the way that submitted assets are evaluated, for example to determine the "maturity" including community review.

E.2.2 CEN/ISSS Workshop eGov-Share

The CEN/ISSS eGov-Share Workshop[3] (more formally the Workshop on Discovery of and Access to eGovernment Resources) was established in February 2008 with the aim to help designers and developers of eGovernment systems and applications by developing approaches and tools to facilitate the sharing of information across agencies and across borders.

As part of a framework that enables the exchange of information between eGovernment agencies, the Workshop defines a reference ontology for eGovernment resources.

The public documentation of OWMS is being used in the workshop as an example of a practical metadata approach that can be mapped to the reference ontology. Further, direct, involvement of the OWMS design and implementation team could help in establishing OWMS as an example for other European countries.

E.3 Global initiative: DCMI

The global Dublin Core Metadata Initiative[4] is also involved in bringing together representatives of eGovernment programmes in the DCMI Government Community and

[1] http://www.semic.eu/semic/view/snav/About_SEMIC/Launch-Conference.xhtml
[2] http://www.semic.eu/semic/view/documents/vision-of-the-clearing-process.pdf
[3] http://www.cen.eu/isss/eGovShare
[4] http://dublincore.org/

DCMI Government Application Profile Task Group[1]. OWMS can continue to play an important role in the global level, potentially as a recognized "showcase" of the use of Dublin Core metadata in the government domain.

E.4 International conferences

E.4.1 ESIIG

The Second European Summit on Interoperability in the iGovernment, ESIIG2[2], that took place in Rome, 20-22 October 2008, adopted the Declaration of Rome[3] proposed the constitution of a European Interoperability Agency as a tool for the promotion of interoperability amongst European administrations, and proposed the creation of ERNI (European Regional Network for Interoperability) as a permanent governance structure among European regions, with the aim of creating a group of preferential interest that plays a role of qualified interlocutor towards the national and European authorities competent in the field.

Whenever further conferences are announced, it could be beneficial for the OWMS team to participate to present OWMS and exchange experience with other implementers of metadata solutions in the government domain.

E.4.2 DEXA EGOV

DEXA EGOV2009[4], the eighth international EGOV conference 2009 within the DEXA conference cluster will take place in Linz, Austria, 30 August-2 September 2009.

The international EGOV conference series gives annual state-of-the-art overviews in eGovernment and eGovernance research, implementation and application. The EGOV conferences bring together researchers and professionals from many countries and many disciplines. In recent years, the assessment of eGovernment efforts, the prospects of eGovernment as a research discipline, and the role of information and communication technology for development rank among the top topics on the research agenda.

Participation of the OWMS to present the implementation and to get insight in current research in issues of interest for future developments may be beneficial.

E.4.3 International Conference on Dublin Core and Metadata Applications

A final opportunity for presentation of OWMS and discussion with other eGovernment programmes around the world is DC-2009, the conference organized by DCMI and the National Library of Korea in Seoul, Korea, 12-16 October 2009.

OWMS could be presented as a project report in the conference track and as input to a possible DCMI Government Community meeting. This could allow a two-way process of exchange of experiences with implementers of Dublin Core-based solutions in other countries.

[1] http://dublincore.org/groups/government/ and http://dublincore.org/dcgapwiki/
[2] http://www.esiig2.it/esiig2/index.php
[3] http://www.esiig2.it/esiig2/doc/Dichiarazione_di_Roma_en.pdf
[4] http://www.egov-conference.org/egov-2009