A comparison of different semantic methods for integrating thematic geographical information: the example of land cover

Alexis Comber[1], Andy Lear[2], Richard Wadsworth[3]
[1]Department of Geography, University of Leicester, LE1 7RH, UK, ajc36@le.ac.uk
[2]Leicester and Rutland Wildlife Trust, Leicester, LE5 2JJ, UK, alear@lrwt.org.uk
[3]Centre for Ecology and Hydrology, Lancaster, LA1 4AP, UK, rawad@ceh.ac.uk

INTRODUCTION

There is a generic problem when analysing multiple thematic datasets that spatial data of the same phenomenon will vary. Such variation is problematic for analysis that seek to integrate multi-temporal data such as land cover, in order to identify the extent and location of any land cover change, and for models that require land cover as part of their input (e.g. climate change models). This paper compares and evaluates the effectiveness of different methods for integrating land cover data based on analysis of their semantics.

Land cover data may vary for a number of reasons (Comber et al, 2005):
– **Technological**: the use of different measuring instruments. For land cover this includes different sensors with different resolving power and different hyperspectral characteristics and other data capture devices.
– **Politics**: the objectives of studies change due to the changing political climate resulting in different data reporting objectives. This can be seen in the shift from vegetation communities in land cover studies to habitats as a result of the Rio Earth Summit, Kyoto, etc.
–**Scientific understanding**: develops as part of the process of scientific enquiry and investigation. In land cover we see the shift in the norm from Boolean pixel based approaches to fuzzy and object based approaches.
Because of these aspects, observed variation between data may have nothing changes in the phenomenon (i.e. land cover) being measured but are as a result of different specifications and conceptualisations of land cover, as manifested in the thematic data. A number of methods have been suggested to overcome the variations in thematic land cover data and to integrate discordant the data through consideration of their semantics. These are introduced in brief below.

1. Translation to a common framework such as the FAO Land Cover Classification System (LCCS) (Di Gregorio and Jansen, 2000). This approach develops a hybrid ontology approach by providing a set of characteristics to describe land cover classes. It relies on matching the features of different land cover classes with LCCS classifiers to model the similarities between classes. A standard set of diagnostic criteria are evaluated by an expert and the encoded in a hierarchical sequence. These can be compared using a set-based similarity matching procedure that determines the ratio of similar attributes in any two classes from the total number of attributes. Approaches like the LCCS define a shared vocabulary, which makes up the basic building blocks of the domain and are thereby used to compare categories. Other examples include the work of Lutz and Klien (2006) and Kavouras and Kokla (2002). The underlying assumption behind these approaches is the characteristics or features or vocabulary describe all of the relevant data characteristics. The advantages of this approach are, particularly the LCCS, is that it has a very high uptake amongst global land cover researchers who have promoted this as 'the standard' for land cover.

2. Semantic consistency, such as suggested by Comber et al (2004). This method uses an expert approach to generate measures of belief in each class pair by scoring the relations between individual class pairs in terms of their relation being consistent, inconsistent or uncertain. As well as this

external measure of consistency, an internal measure can also be generated based on the fuzzy membership of the object (parcel or pixel). This approach explicitly allows for uncertainty by extending aggregation from many-to-one' relationships to 'many-to-many' relationships (i.e. allowing a shift from 'yes/no' to 'yes/don't know/no'). The assumption is that an expert exists who is able to understand how different datasets relate to each other. The advantages of this method are that it generates a spatial distribution of uncertainty and accommodates the relative nature of thematic data.

3. Probabilistic Latent Semantic Analysis using a text mining of data semantics (Wadsworth et al, 2008a). It is based on Latent Dirichlet Analysis (LDA) as a method for quantifying semantic differences between geographic categories. Latent Analysis assumes that underlying and unobserved variables (the latent variables) exist and that they can be used to explain an observed pattern. In Latent Semantic Analysis the pattern is the frequency of words in documents and the latent variables are concepts (ideas) described in the documents. This approach provides a 'bottom up' approach to interoperability problems (*cf* formal ontologies which are 'top down'). The advantage of this approach is that it uncovers the relationships between the different classes and their textual descriptions and indicates the latent concepts that can explain the distribution of words in classes. In this way the results reveal 'hidden' or not easily discernable data concepts.

4. Data primitives or conceptual overlaps (Ahlqvist, 2004) are defined as those dimensions or measurements that describe the processes under investigation at the most fundamental level. They provide information about the building blocks that underpin the concepts of the phenomenon – what they mean and what they represent. Data primitive have been referred as 'dimensions' 'conceptual spaces', 'approximation spaces' and 'domains'. Land cover data primitives have been addressed by a number of workers. Alqvist (2004) calculated the degree of overlap and conceptual distance between land cover classes using four 'approximation spaces'. Comber (2008) identified a number of conceptual dimensions to separate the concepts of land use and land cover. Wadsworth et al. (2008b) identified quantified the conceptual overlap between three Siberian land cover datasets and Comber 2008 applied this method to separate land cover from land use.

METHODS

Case study

The different approaches were applied to a problem that sought to integrate national land cover data from satellite imagery with local habitat data collected by field survey. The objective was to explore the suitability of a number of established (but not widely applied) data integration frameworks based around data semantics in order to support divergent local, regional and national objectives. Regional biodiversity partnerships are responsible for regional habitat reporting to national agencies. However there is considerable within- and between-region variation in all aspects of data collection, data management (e.g. metadata). Local habitat data may be temporally and thematically variable. It is collected by different organisations, often using different habitat classifications, to support varying local priorities and objectives (Gallagher and Calder, 2007; NE and TWT, 2009). The variation in habitat data collection and recording has implications for a number of habitat related activities: it reduces the overall quality and consistency of habitat reporting and is problematic for regional objectives such as identifying habitat opportunities (NE and TWT, 2009).

Land cover classes

The results of applying the different approaches to a local case study are described and evaluated in terms of supporting consistent reporting and change analysis. The analysis will apply the different approaches to generate measures of overlap between 3 local habitat classes (Broadleaved and wet woodland, Mixed Grass, Acid Grass) and 3 national broad habitats (Broadleaved, mixed and yew woodland, Improved Grassland, Neutral Grassland). These are described below.

Local Classes

Broad-leaved woodland & wet woodland

Woodland is an area with almost continuous tree and shrub cover, although grassy rides, ponds and buildings etc. may be present. In Leicestershire and Rutland, woodland is a rare habitat extending over about 4% of the counties. Only 1% is ancient woodland and a substantial proportion of that has been damaged by planting. All ancient woodlands are important because of their rarity and many plants and animals are confined or nearly so to ancient woodlands, including replanted ones. Large semi-natural secondary woodlands are also rare and therefore important. The Leicester, Leicestershire and Rutland Biodiversity Action Plan identifies wet woodland and broad-leaved woodland as priority habitats for action.

Primary criteria: The wood meets one of the descriptions listed below:

Description	Size threshold
included in Leicestershire Inventory of Ancient Woodland	None
with at least 4 species from Ancient Woodland Indicator List Z1 which are Occasional, Frequent, Abundant or Dominant	≥2ha
naturally regenerated	≥ 5 ha
dominated by willow and/or alder with the water table seasonally near or above the surface	≥ 0.25 ha
contains colonies of *Hyacinthoides non-scripta* (native bluebells) ≥ 500m^2	≥2ha

Secondary criteria: The site contains blocks of semi-natural woodland totalling one hectare or more in extent. Woodland sites ≥1ha qualify where adjacent to an existing LWS.

Acid Grass and Mixed Grass

Changes in agricultural practices have severely reduced the herb-rich grassland that was once widespread in Leicestershire and Rutland. Consequently, calcareous and neutral grassland are both listed as priority habitats within the Leicestershire and Rutland Biodiversity Action Plan, while acid grassland is covered by the Heath Grassland Action Plan. With the demise of agricultural grasslands, roadside verges have become important refuges for some plant species and these are covered by the Roadside Verge Action Plan. Herb-rich grassland is a fragile habitat and, in most cases, impossible to recreate or restore once it has been damaged. Consequently, the protection of grassland is an essential component of any nature conservation strategy. There are several types of grassland represented in Leicestershire and Rutland, but some have survived agricultural changes better than others and so require different site size thresholds. Grass verges often support the only herb-rich grassland left in an area. These linear grasslands are often small in extent - less than 10 metres wide - but are rich in species. Roadside verges are particularly important for calcareous grassland. Many have been designated as roadside verge nature reserves.

Acid grassland in Leicestershire is naturally poor in plant species. It may occur in association with bare ground and rock outcrops. It may also contain heather and other ericaceous shrubs and can grade into heathland. Although invasion by bracken has led to losses of this habitat, it can survive amongst bracken.

Mixed grassland may support a mosaic of mesotrophic, wet, acid and calcareous grassland types depending on the underlying substrate, hydrology, aspect and other physical features of the site. Quarries, spoil tips, railways and other post-industrial sites often support mixed grassland habitats of great diversity.

National Classes

Neutral Grassland
This broad habitat type is characterised by vegetation dominated by grasses and herbs on a range of neutral soils usually with a pH of between 4.5 and 6.5. It includes enclosed dry hay meadows and pastures, together with a range of grasslands which are periodically inundated with water or permanently moist. Neutral grasslands are sometimes referred to as mesotrophic grasslands. The plant species assemblages that develop on neutral soils are different from those that develop on acid soils (acid or calcifugous grassland) and calcareous soils (calcareous or calcicolous grassland). For the most part neutral grassland communities have few diagnostic indicator species but lack strong calcicoles or calcifuges characteristic of base-rich and acid soils respectively. The National Vegetation Classification describes 12 types of unimproved and semi-improved neutral grassland (Rodwell 1992).
MG1: Arrhenatherumelatius grassland
MG2: Arrhenatherumelatius-Filipendulaulmaria tall-herb grassland
MG3: Anthoxanthumodoratum-Geranium sylvaticum grassland
MG4: Alopecuruspratensis-Sanguisorbaofficinalis grassland
MG5: Cynosuruscristatus-Centaureanigra grassland
MG6: Loliumperenne-Cynosuruscristatus grassland (part only)
MG8: Cynosuruscristatus-Calthapalustris grassland
MG9: Holcuslanatus-Deschampsiacespitosa grassland
MG10: Holcuslanatus-Juncuseffusus rush pasture
MG11: Festucarubra-Agrostisstolonifera-Potentillaanserina grassland
MG12: Festucaarundinacea grassland
MG13: Agrostisstolonifera-Alopecurusgeniculatus grassland
Unimproved or species-rich neutral grasslands are usually managed traditionally as hay-meadows and pastures. Semi-improved neutral grasslands are also included in this broad habitat type and these grasslands are usually managed for pasture or for silage or hay. Neutral grassland differs from improved grasslands by having a less lush sward, a greater range and higher cover of herbs, and usually less than 25% cover of perennial rye-grassLoliumperenne.

Improved Grassland
This broad habitat type is characterised by vegetation dominated by a few fast-growing grasses on fertile, neutral soils. It is frequently characterised by an abundance of rye-grassLolium spp. and white clover Trifoliumrepens. Improved grasslands are typically either managed as pasture or mown regularly for silage production or in non-agricultural contexts for recreation and amenity purposes; they are often periodically resown and are maintained by fertiliser treatment and weed control. They may also be temporary and sown as part of the rotation of arable crops but they are only included in this broad habitat type if they are more than one year old. Sown grasslands which are less than one year old are included in the 'Arable and horticultural' broad habitat type.

Broadleaved, mixed and yew woodland
This broad habitat type is characterised by vegetation dominated by trees that are more than 5 m high when mature, which form a distinct, although sometimes open canopy with a canopy cover of greater than 20%. It includes stands of both native and non-native broadleaved tree species and yew Taxusbaccata, where the percentage cover of these trees in the stand exceeds 20% of the total cover of the trees present. Woodlands that are dominated by conifer trees with less than 20% of the total cover provided by broadleaved or yew trees are included in the 'Coniferous woodland' broad habitat type. Stands of broadleaved, mixed and yew woodland may be either ancient or recent woodland and either semi-natural arising from natural regeneration of trees, or planted. Recently felled broadleaved, mixed and yew woodland is also included in this broad habitat type where there is a clear indication that it will return to woodland. Otherwise it is classified according to the field layer composition. Scrub vegetation, where the woody component tends to be mainly shrubs usually less than 5 m high, and carr (woody vegetation on fens and bog margins) is included in this broad habitat type if the woody species form a canopy cover of greater than 30% and the patch size of scrub is greater than 0.25ha. Exceptions to this include dwarf gorse Ulex minor and western gorse Ulexgallii which are included in the 'Dwarf shrub heath' broad habitat type, montane willow scrub which is included in the 'Montane habitats' broad habitat type, and scrub on sand dunes and shingle which is included in 'Supralittoral sediment' broad habitat type. Stands of bog-myrtle Myrica gale are included in this broad habitat type as scrub if they are more than 1.5 m tall. This habitat type does not include hedges (woody vegetation that has been managed as a linear feature) as these are included in the 'Boundary and linear features' broad habitat type.

RESULTS

1. Translation to a common framework

To calculate the overlaps between classes the LCCS compares the two LCCS codes that are generated as a result of the LCCS classification process (Table 1). For the evaluation and feature matching one class is selected as the referent class and then a set-based matching process calculates the ratio of similar attributes between the two classes (Table 2). Ahlqvist (2008) noted some of the limitations of the LCCS: The LCCS classifiers may not match class definitions; some classes are described by few classifiers; feature matching assumes equal salience for all features

Local	Acid Grass	A6A10B4C1-N1101
Local	Mixed Grass	A6A10B4C2
Local	Broadleaved woodland	A3A10B2C1D1E2
National	Neutral Grassland	A6A10B4C1
National	Improved grassland	A4-xx-B5C1D1
National	Broadleaved, mixed and yew woodland	A3A10B2-xx-D1E2F2F6F7G3

Table 1. Class codes generated by the LCCS classification process

To / **From**	Acid Grass	Mixed Grass	Broadleaved woodland	Neutral Grassland	Improved grassland	Broadleaved, mixed and yew woodland
Acid Grass		0.75	0.25	**1.00**	0.25	0.25
Mixed Grass	**0.75**		0.25	**0.75**	0.00	0.25
Broadleaved woodland	0.33	0.17		0.33	0.33	**0.83**

Neutral Grassland	**1.00**	0.75	0.50		0.25	0.25
Improved grassland	0.25	0.00	0.50	0.25		0.25
Broadleaved, mixed and yew woodland	0.11	0.11	0.56	0.11	0.11	

Table 2. Measures of semantic overlap using LCCS feature matching (>= 0.75 in bold)

2. Evaluations of internal and external data consistency

An expert identifies the "consistent" (+1), "inconsistent" (-1) and "uncertain" (0) relations between classes (Table 3). These values can be used to generate measures of belief in consistency between datasets when data are overlaid and if sub-object information is available (pixel membership functions or parcel composition) measures of internal consistency. It is problematic when more than two classifications need to be compared die to the high number of pair-wise comparisons that the expert needs to make, and experts are notoriously unreliable – they vary, change their mind and cannot always explain their decisions.

National **Local**	Neutral Grassland	Improved grassland	Broadleaved, mixed and yew woodland
Acid Grass	1	0	-1
Mixed Grass	1	0	-1
Broadleaved woodland	-1	-1	1

Table 3. A semantic look up table of consistency relations between classes

3. Probabilistic Latent Semantic Analysis

The similarity of the classes was examined as in Wadsworth et al (2006) using Probabilistic Latent Semantic Analysis (PLSA). This was proposed by Hofmann (1999a,b) as a "generative" model of latent analysis; the joint probability that a word (w) and document (d) co- occur (P(d,w)) is a function of two conditional probabilities; that the document contains a concept (z) (P(z|d)) and that the word is associated with that concept (P(w|z) (Equation 1)

$$P(d,w) = P(d)\sum_{z \in Z} P(w \mid z)P(z \mid d) \qquad \text{(Equation 1)}$$

Using the frequency of the words in documents (*n(d,w)*) it is possible to rearrange the probabilities to develop an iterative expectation maximization scheme to estimate all the probabilities. The expectation step generates P(z|d,w) while the maximization step calculates P(w|z), P(d|z) and P(z). The distances in semantic feature space can be plotted from the PCA (i.e. the axes do not relate to a specific attribute) and the distances between clases can be visualised in this PCA space.

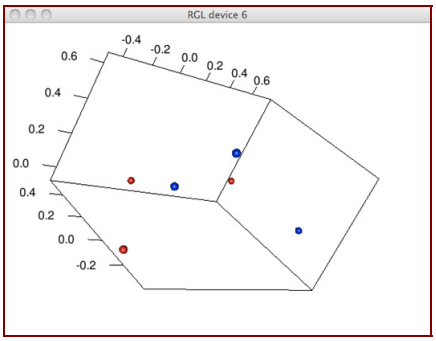

Figure 1. Plot of the PCA of the 3 local classes (blue) and the 3 national classes (red)

Additionally the words associated with each of the 12 dimensions or topics identified by the PLSA on an anlysis of the full text of the 15 local habitat classes and the 17 national broad habitat classes can be plotted.. Figure 2 shows the concepts associated with each of the PLSA topics and Figure 3 shows how each the links between these and the different classes from the two classification schema.

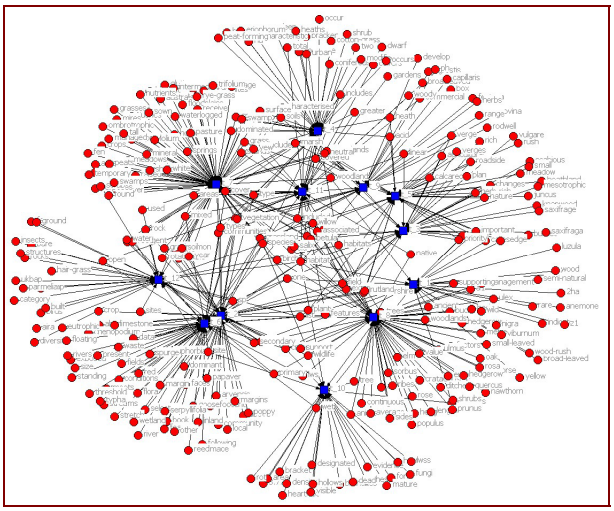

Figure 2. The topics identified by PLSA (blue) and the terms associated with each of these

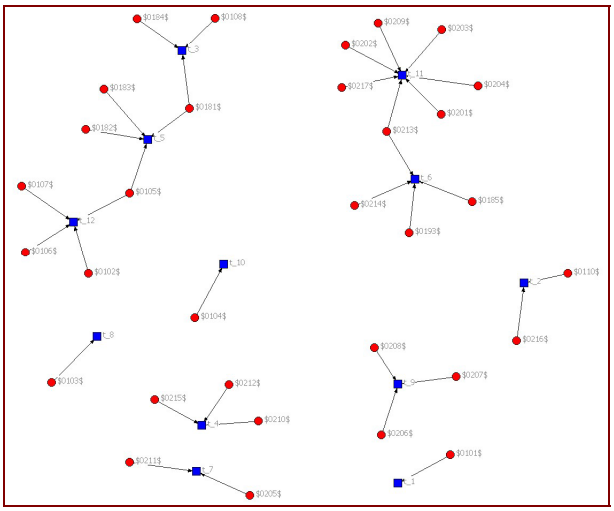

Figure 3. The classes (red) associated with each of the PLSA terms (blue). The following labels indicate the examples in this short paper: 0101 Local Broadleaved woodland; 0183 Local Acid grass; 0185 Local Mixed grass; 0201 National Broadleaved woodland; 0205 National Improved grassland; 0206 National Neutral grassland

4. Data primitives or conceptual overlaps

Six primitive dimensions were identified and scored by the proposing team in order to illustrate the method. We note that in the actual project this will be done in collaboration with all regional biodiversity / habitat reporting partners. The dimensions were:

- vegetation structure or canopy cover (Complex to Simple);
- biomass harvesting (Most to Least);
- vegetation height (High to Low);
- soil fertility / agricultural improvement (Most to Least);
- species richness (Most to Least);
- size / area (Large to Small).

The dimensions were all scored in a non-ordered qualitative manner using 10 classes, although we note that some could be continuous (e.g. vegetation height, biomass harvesting, area etc). Each class was scored as being present (1) or absent (0) and Equation 2 was applied to them. The example of Species Richness is shown in Table 4.

Type	Class	High									Low
Local	Acid Grass								1	1	1
Local	Mixed Grass		1	1	1	1					
Local	Broadleaved Woodland					1	1	1			
National	Improved Grassland									1	1
National	Neutral Grassland	1	1	1	1	1	1	1	1		
National	Broadleaved, mixed & yew woodland		1	1	1	1	1	1	1	1	

Table 4. Scoring of species richness

Each class can be assessed more or less independently within each domain. We note that complete independence is not possible for qualitative domains such as 'biodiversity value' because

the allocated value for any class in that domain is relative to the other classes. Measures from Bouchon-Meunier et al (1996) can be applied to both continuous domains such as 'canopy cover' or tree height and to non-ordered qualitative domains as in this case (Equation 2):

$$O(p_A, p_B) = \frac{\sum \min(p_A, p_B)}{\sum (p_B)} \qquad \text{(Equation 2)}$$

where $f_{pA}(x)$ and $f_{pB}(x)$ represent the values of concept (classes) A and B at location x in domain p; and pA and pB are the properties of concepts A and B in domain p. The overlap measure can vary from 0 (no overlap) to 1 (class B is a subset of A). Classes will overlap to a different degree in each of the domains. The overlaps between the different classes are shown in Table 5.

		Local			National		
	To / **From**	Acid Grass	Mixed Grass	Broadleaved Woodland	Improved Grassland	Neutral Grassland	Broadleaved, Mixed and Yew woodland
Local	Acid Grass		0.72	0.06	0.40	0.60	0.25
	Mixed Grass	0.56		0.25	0.28	**0.85**	0.47
	Broadleaved woodland	0.02	0.22		0.14	0.31	**0.83**
National	Improved Grassland	0.44	0.28	0.19		0.32	0.25
	Neutral Grassland	0.51	**0.82**	0.25	0.33		0.43
	Broadleaved, Mixed and Yew woodland	0.12	0.32	0.72	0.22	0.35	

Table 5. The conceptual overlaps between local and national classes

DISCUSSION

The results of the four approaches for integrating data semantics show a number of characteristics. Two of the methods, 'Translation to a common framework' and 'Data primitives or conceptual overlaps' generate measures of correspondence between the classifications (Tables 2 and 5 respectively). Although these are generated by different processes, both rely on some form of subjective interpretation of class characteristics, for instance by experts. Neither of the sets of correspondences is symmetrical. They can be applied to give measures of similarity between the two classes (and in turn the two classifications) and can be used to develop alternatives mappings of landscape features for each class of interest. This is in a manner similar to fuzzy memberships, generating a layer for each new class. The correspondence measures derived from are applied consistently across all objects, with no account of any local variations. The 'Expert evaluations of internal and external data consistency' method generates a much simpler table of relations between classes using a three-valued logic to indicate consistency and uncertainty. These measures can be

used to provide more detailed object specific measures of correspondence where the data have internal measures of consistency such as are commonly generated by object based image analysis from remotely sensed data. That is, the measures of correspondence vary locally depending on the characteristics of the data object they are applied to. Lastly, 'Probabilistic latent semantic analysis' mines the data descriptions to identify an optimum number of latent variables or 'topics', in this case 12 (Figure 2). These describe distinctive 'themes' in the data. In turn the classes that are most strongly linked to each topic can be identified (Figure 3). This approaches replaces the need for the (human) expert by identifying the semantic concepts that are associated with each class and thereby the topics that are shared between different classes from different classifications. This bottom up approach has the potential to support automated methods for data integration by automatically identifying semantic overlaps. However, further work is needed to develop methods to interpret the topics we note that one of the problems with the approach is that repeated "runs" on the same data set do not always result in the same topics being identified.

CONCLUDING REMARKS

Each of the approaches to identifying semantic similarities and differences described provide some measure of overlap between different classes. The approaches apply very different techniques and accommodate to differing degrees some of the uncertainty associated of moving from one classification to another. We note that Probabilistic latent semantic analysis provides a method for automated mining of data semantics, where the dimensions may not be known *a priori*, and that such 'bottom up' approaches to integrating data semantics have the potential to support eScience infrastructures such as INSPIRE. Future work will apply the overlaps to land cover and habitat data.

REFERENCES

Ahlqvist, O. (2004). A Parameterized Representation of Uncertain Conceptual Spaces. *Transactions in GIS*, 8: 493–514.

Ahlqvist, O. (2008). In search of classification that supports the dynamics of science: the FAO Land Cover Classification System and proposed modifications. Environment and Planning B: Planning and Design 35: 169-186.

Bouchon-Meunier, B., Rifqi, M. and Bothorel, S. (1996). Towards general measures of comparison objects. *Fuzzy sets and systems*, 84: 143-153.

Comber, A., Fisher, P., Wadsworth, R., (2004). Integrating land cover data with different ontologies: identifying change from inconsistency. *International Journal of Geographical Information Science*, 18(7): 691-708.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., (2005). What is land cover? *Environment and Planning B: Planning and Design*, 32:199-209.

Comber, A., (2008).The separation of land cover from land use with data primitives. *Journal of Land use Science*, 3(4): 215–229.

Di Gregorio, A., and Jansen, L.J.M. (2000). *Land Cover Classification System: Classification Concepts and User Manual*, Rome: FAO.

Gallagher, P. and Calder, G. (2007). Biodiversity of Scottish Wildlife Trust Reserves http://www.swt.org.uk/docs/002_001_publications_BiodiversitySWTReserves07_1250 595197.pdf [Available 17/08/09]

Hofmann, T. (1999a). Probabilistic latent semantic indexing, pp 50-57 in *Proceedings of 22nd International Conference on Research and Development in Information Retrieval* (Eds Hearst M, Gey F, Tong R) Univ Ca, Berkeley, California, Aug, 1999

Hofmann, T. (1999b). Probabilistic latent semantic analysis, pp 289-296 *in Proceedings of 15th Conference on Uncertainty in Artificial Intelligence* (Eds. Laskey KB, Prade H) Royal Inst Technol, Stockholm, Sweden, Jul 30-Aug 01, 1999

Kavouras M. and Kokla M. (2002). A method for the formalization and integration of geographical categorizations. *International Journal of Geographical Information Science*,16: 439-453

Lutz M. and Klien E., (2006). Ontology-based retrieval of geographic information. *International Journal of Geographical Information Science* 20: 233-260

NE and TWT (Natural England and the Wildlife Trusts) (2009). *6Cs Growth Point Biodiversity Opportunity Mapping*.http://www.emgin.co.uk/default.asp?PageID=261 [available 17/08/09]

Wadsworth R.A, Comber A.J., and Fisher P.F., (2006). Expert knowledge and embedded knowledge: or why long rambling class descriptions are useful. pp 197 – 213 in *Progress in Spatial Data Handling, Proceedings of SDH 2006*, (eds. Andreas Riedl, Wolfgang Kainz, Gregory Elmes), Springer Berlin.

Wadsworth, R.A, Comber, A.J. and Fisher, P.F., (2008a). Probabilistic Latent Semantic Analysis as a potential method for integrating spatial data concepts, pp 99-108 in *Proceedings of the Colloquium for Andrew U. Frank's 60th Birthday*, (ed. Gerhard Navratil), GeoInfo Series 39, Vienna, ISBN 978-3-901716-41-6

Wadsworth, R., Balzter, H., Gerrard, F., George, C., Comber, A.J., and Fisher, P.F. (2008b). An Environmental Assessment of Land Cover and Land Use Change in Central Siberia Using Quantified Conceptual Overlaps to Reconcile Inconsistent Data Sets. *Journal of Land Use Science*, 3(4): 251.