MINISTRY OF DEFENCE

Classification & Reference Data Support for MODAF

Version 1.0

12 April 2006

Prepared by:- Ian Bailey, IA8bCon5

Approved by:- MODAF Technical Working Group

RECORD OF CHANGES

This page will be updated and re-issued with each amendment. It provides an authorisation for the amendment and a checklist to the current amendment number.

Issue No.	Date	Revision Details
Draft 0.1	28 January 2006	First Draft for Comments
Draft 0.2	12 March 2006	Second Draft following review by Peter Bryant
Draft 0.3	20 March 2006	Minor changes following further internal review.
Release 1.0	12 April 2006	Significant changes following comments from CDMA. Most significant being the change in name from "MODAF Taxonomy" to "MODAF Ontology".

Table of Contents

Summary

Readers may have been expecting references to "The MODAF Taxonomy", a term that has been used consistently across all previous MODAF documentation. The need for a standard classification and reference facility for MODAF was recognised early in its development and the associated requirements endorsed by the MODAF Technical Working Group. DG-Info (CDMA) had the responsibility for delivery of what was initially called "The MODAF Taxonomy". In the current phase of work, CDMA has highlighted that the term "Taxonomy" is not appropriate, as the formal definition of a taxonomy, as used by MOD, is a set of terms which are structured into broader and narrower terms whereas the MODAF requirements (see App B) specify something more sophisticated than this. Accordingly CDMA has directed that "Structured Library" or "Ontology" would be more appropriate. As a result, what was to be called the "MODAF Taxonomy" is now to be called the "MODAF Ontology" (this change is subject to ratification by the MODAF Technical Working Group). Now that the MODAF baseline documentation set and Meta-Model (see sub-sections below) have been delivered, the MODAF Ontology will have to be delivered as part of on-going MODAF maintenance. At a future date, CDMA will assume ownership of the MODAF Ontology as a component of its "Information Layer".

This document reports the findings of an analysis of applicability of the UK Defence Taxonomy (UKDT) and the IDEAS Model to MODAF. In brief, the UKDT is extensive and well-thought-out. It is structured for search and document classification rather than for architectural support, and it is this issue which exercised the majority of the analysis presented in this document.

The UKDT contains much information that is useful to MODAF architects, and would provide an excellent (if not deep) coverage of most architectural elements. MODAF would require that the UKDT be significantly extended, for example, in the area of technical systems, operational activities, platforms and information formats. When this is considered in light of the fact that the UKDT's structure is not suitable for MODAF, the conclusion is that UKDT contains much that can be re-used, but is not useful to MODAF in its current form.

The key issue is one of management. Given that the structure of the UKDT is not suitable for MODAF use, but that much of the content is, an approach to managing an alternative structure is required. As MODAF is likely to extend the UKDT in the technical subject areas, it is thought best that a MODAF Ontology is established which re-uses the UKDT terms by means of a formal reference (e.g. Dublin Core or MOD Meta-Data Standard). This will inevitably involve regular maintenance to avoid the taxonomies getting out of synchronisation. In the long term, there is a will within CDMA to manage the MODAF Ontology. However, until the CDMA Information Layer is in place, there is a requirement for an interim governance framework.

In addition, this report examines the IDEAS model. There is a will within the MODAF community to adopt the IDEAS model as a replacement for the ERM should it be suitable. As well as the benefit of the IDEAS model being a "from-scratch" design, there is a significant advantage to having a model (and perhaps the upper levels of an ontology) that is common between coalition partners. This report attempts to document MOD's requirements for IDEAS, from the point of view of IDEAS being used as the top-level of the MODAF Ontology. To meet this requirement, the UK must ensure that IDEAS follows a rigorous, ontological approach.

This document also provides a background primer on taxonomy and ontology, the UK Defence Taxonomy (UKDT) and the IDEAS model.

1 Introduction

The purpose of this document is to report initial findings on the applicability of the UK Defence Taxonomy and the IDEAS[1] model to MODAF. The analysis reported in this document took place after the MODAF Technical Working Group approved an outline plan to investigate the possible replacement of the current ERM with the IDEAS model, and whether the UKDT was suitable for extending the concepts in the ERM/IDEAS model. With the new name of "MODAF Ontology", the plan would be for the IDEAS model, if suitable, to act as an "upper ontology" defining the core concepts and their relationships. The UKDT would then be used to continue specialising those concepts into a "lower ontology" (what was originally to be called the "MODAF Taxonomy").

The importance of re-using existing MOD taxonomies and glossaries was recognised early on in the MODAF project – the task of developing an extensive taxonomy from scratch is labour-intensive. As one would expect for an organization the size of the MOD, there are many sources of glossary, taxonomy and reference data. There is, however, only one centrally standardised source for such data – the UK Defence Taxonomy (with its associated thesaurus). The two questions which drove this investigation were;

- How suitable is the UK Defence Taxonomy for use with MODAF, and what effort is required to make it suitable ?
- Is the IDEAS Model likely to be suitable as an upper ontology for MODAF ?

1.1 The MODAF Enablers

The MOD Architectural Framework (MODAF) is underpinned by 4 main technical enablers; the MODAF Meta-Model[2] (M3), the Enterprise Reference Model[3] (or ERM, which may soon be re-worked as the MODAF Ontology), the newly renamed MODAF Ontology[4], and the repository (MODAR). The M3 defines a UML profile, which has the main purpose of enabling meaningful data exchange using the XMI[5] specification, though it may also be used to configure the meta-models of UML™ 2.0 tools. In addition to the three information standards, there is the repository implementation– MODAR[6]. This will specify a standard repository for MODAF-compliant architectural data.

[1] www.ideasgroup.org – The UK MOD is represented through DG-Info

[2] See IA document: IA/02/16-ERMcm04

[3] See IA document: IA/02/16-ERMcm06

[4] See IA document: IA/02/16-ERMcm05

[5] "XMI" stands for XML Meta-Data Interchange, and is the OMG™ standard for exchanging models. For the initial MODAF release, XMI v2.1 for UML™ v2.0 is to be used.

[6] At the time of writing, the final specification for MODAR has yet to be released. Contact the Integration Authority for updates

1.2 The MODAF Enterprise Reference Model (ERM)

The ERM is a conceptual model that defines the requirements for MODAF architectures. The scope of the ERM is to cover anything that relates to MODAF. As MODAF is an enterprise architecture, the scope is very broad, including such aspects as:

- Operational and business processes
- People, Organizations and their skills
- Operational an business constraints
- Projects
- Capabilities (required and actual)
- Systems & Materiel

The ERM pre-dates MODAF and is the merger of four pre-existing data models – each with its own stakeholders. The intention was that the ERM would cover the scope of all these data models, by merging the models into one. However, the models were defined at different levels of abstraction (logical and conceptual) and were oriented differently (some define the architectural elements, others the real world objects that the architectures represent). In addition, the stakeholders were keen that their ideas made it into the final ERM – hence there was redundancy in the model.

It was immediately obvious the ERM was not suitable to act as an implementation model for MODAF. So, to ensure that the MODAF baseline deliverables were complete, the MODAF Meta-Model (M3) was released along with the baseline documents as a draft (v0.95) and was released as v1.0 on 12th April 2006. The M3 is an implementation model. Though it probably could not be described as a physical model, its purpose is to define the structure of XMI exchange files and architectural repositories. Around the time of the MODAF baseline release, the ERM underwent a rationalisation in an attempt to reduce redundancy in the model. Again, this was difficult to achieve without compromising the preferred approach of one or more stakeholders. The resulting ERM (v2.1) was a great improvement, however when mapped to the M3[7], it was apparent that the ERM fell short of the complete MODAF scope. Also, because of the disparate nature of the source models, the resulting ERM did not have a useful inheritance hierarchy – something which would be essential if it were to be used as the upper level of the MODAF Ontology.

The conclusion, therefore, was that the ERM has been useful as a collector of requirements, but was not an accurate representation of MODAF, and was not well suited to being a foundation for the MODAF Ontology. It carried with it too much “baggage” from the models on which it was based. At the MODAF Technical Working Group meeting on 18th November 2005, it was decided that MOD should “freeze” the ERM and replace it with a model designed from scratch that represented MODAF more accurately. In particular, it was decided that the enablers team should investigate the possible use of the IDEAS model as the basis for this Ontology.

---

[7] See IA Document - IA/02/16-ERMcm07

1.3 The IDEAS Model

The aim for the MODAF enablers is to have a model that can act as an upper ontology (see Appendix A for background) to establish a rigorous framework which can be extended by elements from the UKDT and other sources of reference data such as the ISSE Library. As discussed in Section 1.2, the ERM is not appropriate for this, so the IDEAS model is under consideration as a candidate basis for the MODAF Ontology.

The IDEAS model is a conceptual model for military architectures and is being developed by the UK, Australia, Canada and Australia (with NATO as observers). The MODAF Technical Working Group has requested that the UK team contributing to the IDEAS model should try to ensure that the model was

a) suitable for MODAF

b) structured appropriately to act as the upper level for the taxonomy - i.e. it should be as close to a formal Ontology (see Appendix A) as possible.

Also, semantic compatibility with M3 will be a key goal for the UK.

Initial work on the IDEAS model has been focused towards concepts needed to cover a coalition operations scenario from a communications point of view. This has meant the model concentrates on force structure, systems connections, protocols and IERs. The IDEAS Group will meet again in May ’06 to further develop the model, and aim to have complete coverage of DoDAF by the Q1 ’07. MODAF’s scope is greater than DoDAF (because of the additional Strategic and Acquisition views), so the UK will probably have to extend IDEAS if it is to be suitable as a replacement for the ERM (see previous section).

1.4 The MODAF Lower Ontology (previously: “Taxonomy”)

The MODAF Lower Ontology is the main focus of this document, however the ERM/Upper Ontology also has a strong bearing on the subject. The MODAF project requires a standard classification scheme in order to ensure a common use of terminology for the appropriate architectural elements. An architect should be in no doubt about what types of systems, organizations, etc. are at his/her disposal to use. The architect should also be able to understand what types have been used in other MODAF-compliant architectures. To enable this level of architectural commonality and transparency, the lower ontology used must be defined in reference to the types of architectural element defined in the MODAF Meta-Model. For example, where the M3 only goes to the level of defining systems, the taxonomy should provide more in-depth information about what kind of systems are recognised.

It would therefore be ideal to use M3 as the top-level of the ontology. However, because M3 is a meta-model, it defines types of architectural element (i.e. things that appear in models). Most of this information is not useful in an ontology. The MODAF Ontology *is* intended to provide the meaning behind MODAF, without any of the meta-model baggage. The Upper Ontology would define the most general concepts and the relationships between them, and the Lower Ontology would specialise those concepts into more specific classes of system, platform, organization, activity, etc. This is illustrated in Figure 1.

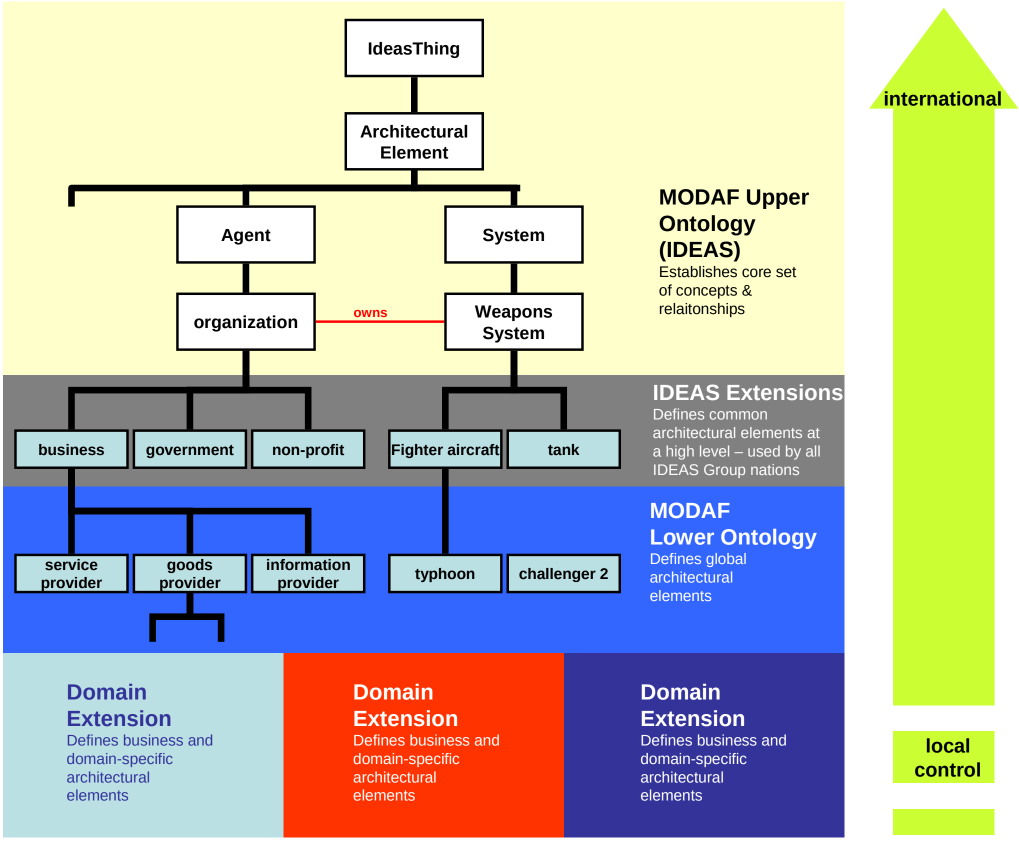

Figure 1 - The MODAF Ontology Layers

Figure 1 also shows how the MODAF Ontology may be extended further by other interested parties – e.g. an IPT adding its own, project-specific elements. It is worth clarifying that the borders between taxonomy levels are not fixed. For example, should an IPT invent new elements that are then useful to other projects, there would be a strong case for formalising those elements into the core MODAF Ontology. Figure 1 also shows an aspiration towards an international taxonomy (labelled "IDEAS Extensions") – i.e. an agreed set of terminology between the coalition nations. This has not been agreed, but the benefits are clear – ability to exchange and share architectural information between nations with at least a basic level of understanding of the types of elements in the architectures.

The requirements for the MODAF Ontology (called "Taxonomy" at the time of writing) have been reproduced in Appendix B of this document.

1.5 The UK Defence Taxonomy & Thesaurus

The UK Defence Taxonomy has been developed to enable file and document classification for search purposes. It has also found application in other areas, such as being the terminology reference for the MOD Business Management System, and the official reference for classifying MOD Estate. The taxonomy has been developed according to a strict sub-classification regime based on narrowing terms, and is a tree structure (single inheritance[8]). The upper-most branches of the taxonomy are subject areas, these are:

- Defence Estate
- Defence Policy and Strategic Planning
- Equipment and Materiel
- Exports Disposals and Sales
- Finance
- Health
- Information Policy and Services
- Legal
- Logistic Support
- Management and Communication
- Operations and Operational Training
- Organizations their Role and History
- Personnel
- Procurement
- Research Science and Technology
- Safety Environment and Fire
- Security and Intelligence
- Support Services
- Training and Education

The coverage of the taxonomy is fairly broad, with the tree really only being 2-3 levels deep for most subject areas. Depth is added by the Defence Thesaurus which extends the tree to deeper levels and introduces synonyms for the approved terms.

[8] This means that a given term can only have one broader term – i.e. it can only descend from one branch of the tree

2 The Analysis

2.1 Analysis Procedure

The initial investigation of the UK Defence Taxonomy involved converting its native XML format to the W3C Web Ontology Language (OWL). This then enabled the data to be viewed and manipulated using freely available tools. As the current (v2.1) ERM model is not suitable to act as an upper ontology for MODAF (see Sections 1.2 and 1.3), it was decided to use the top level concepts from the first draft of the IDEAS model. This model was also converted to the OWL format so that it could be directly compared and contrasted with the UKDT.

The comparison work involved modifying an OWL editing tool to enable the UKDT terms to be "dragged and dropped" under the most appropriate concepts in the IDEAS model. The main effort of the comparison was, therefore, a first attempt at placing the UKDT terms in the correct place under the IDEAS model.

The resulting structure is shown in Appendix C. Note that this was by no means an exhaustive attempt to fully map the UKDT, so should be treated as purely indicative of the approach that will be needed to build the MODAF Taxonomy.

3 Findings

3.1 The UK Defence Taxonomy

The UKDT has been designed primarily to facilitate searching. It is a detailed and precise list of terms, structured by subject area. The taxonomy terms are refined into broader and narrower terms. As a document and search classification taxonomy, it is an excellent piece of work – it is precise, clear and very easy to navigate. It is clear that the UKDT is *the* MOD reference for terminology (with additional support from the thesaurus and the Defence Data Repository), so it would be inappropriate for MODAF not to use it. However, the findings of this study (see following sub-sections) reveal that a certain amount of re-structuring will be required to use the UKDT for MODAF. The key to success here is to find a way that the current UKDT structure can co-exist with a MODAF-oriented version, with all the configuration management issues that arise from maintaining two structures of the same information.

3.1.1 Pure Specialisation

The collaborative architecting requirement for MODAF is subtly different to the search requirements of the UKDT. The semantics of MODAF is captured in the Enterprise Reference Model (or a possible future MODAF Upper Ontology). The ERM/Upper Ontology will define the core classes, with the intention that the MODAF Lower Ontology will specialise those concepts into more domain specific classes. For example, where the Upper Ontology might define "Operational Activity", the Lower Ontology will define specific types of activity such as "Conduct Battle Damage Assessment", "Transport Materiel", etc. Architects developing an activity model would not expect to be able to classify their activities in any other way – e.g. you would not classify your activity as a type of document. It is this requirement for pure specialisation (i.e. all elements in a branch of the lower ontology must be of the same base type – e.g. Operational Activity) that results in the UKDT not being directly applicable to MODAF without modification. The branches of the UKDT specialise by subject area rather than ontological type – i.e. it is possible to find elements of different basic type in the same place. This issue is exemplified in Figure 2, where the finance subject area contains a term "Financial Management" which would be an Operational Activity in MODAF, but which contains policy (a standard in MODAF), services (Operational Activity), training (Operational Activity), Statistics (Information), etc.

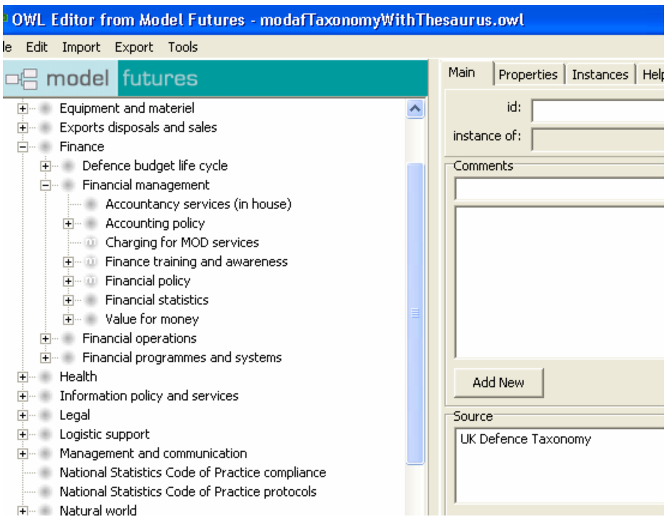

Figure 2 - UKDT specialisation tree for finance

3.1.2 Classes & Individuals

The UKDT does not distinguish between classes and individuals (see Appendix A for an explanation of these terms), it merely defines broader and narrower terms. The case of warships is a perfect example. There are classes of Naval vessels, and there are the vessels themselves. In the UKDT, HMS Daring is just a narrower term of the Type 45 Daring Class vessel (see Figure 3). It is essential that an architecture clearly distinguishes between classes and individuals. In some architectures it may be sufficient to specify just that a Type 45 is used. However, for an architecture for HMS Daring, the architect would want to be clear that the individual was the focus. From MODAF's point of view, therefore, there is a requirement to distinguish between classes and individuals.

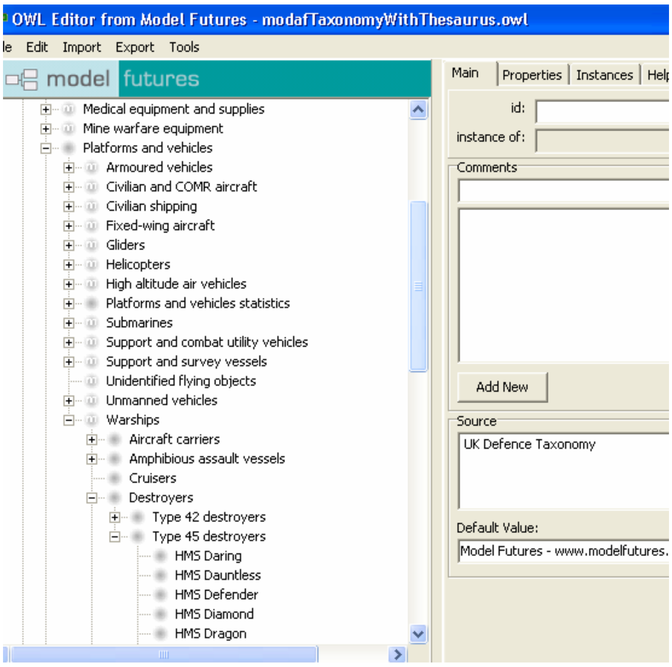

Figure 3 - Type 45 destroyers

3.1.3 Relationships

In a similar fashion to its treatment of classes and individuals, the UKDT tends to treat relationships between individuals (esp. whole-part composition relationships) as broader-to-narrow terms. Again, this is eminently suitable for search purposes, but more precision is needed for MODAF. This example is shown clearly in the case of defence estate where garrisons and barracks contain individual barracks and garrisons – the barracks that make up a garrison are shown as narrower terms when in fact they should be parts of the garrisons (see Figure 4).

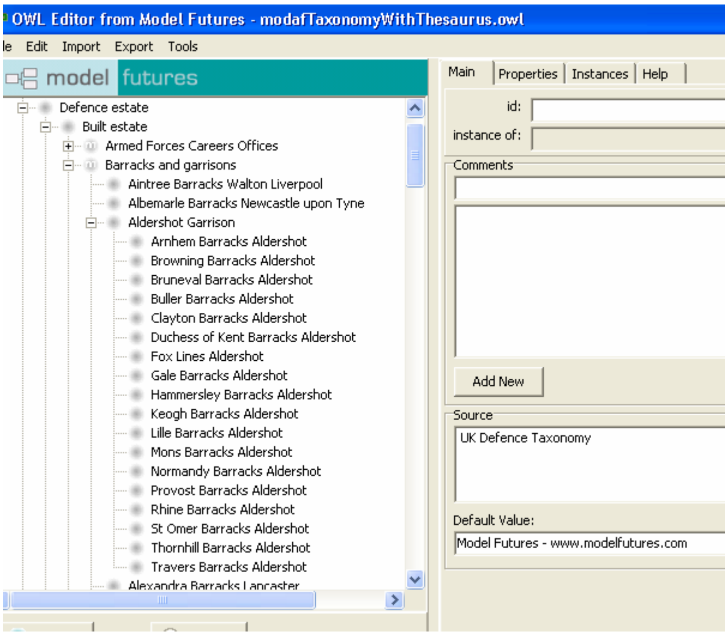

Figure 4 - Barracks & garrisons

3.1.4 UKDT Content

MODAF users will require the greatest depth and detail in the areas of weapons systems, communications systems, platforms and information specifications. Although the UKDT covers these areas, the depth would probably not be sufficient for most architects. For this reason, it is likely that the MODAF Lower Ontology will augment the UKDT in these areas. It is unclear whether the UKDT will require such a level of detail, so these elements may just reside in the MODAF Lower Ontology.

3.2 The IDEAS Model

At the time of conducting the analysis on the UKDT, the IDEAS team had just established the base concepts:

- Activity
- Association
- Information Item
- Mission
- Organization
- Project
- Requirement
- Scenario
- Technology
- Agreement
- Capability
- Location
- Node
- Port
- Property
- Rule
- System
- Unit of Measure

These concepts were sufficient to allow an analysis of the UKDT, however. Since this analysis, in Q1 of 2006, the IDEAS Group met to commence the modelling task. The team will meet again in May 2006 to continue the modelling effort. Following that meeting, there should be sufficient coverage for the MOD to make a judgement on the suitability of IDEAS as the conceptual model / upper ontology for MODAF. The issue facing MOD with the

IDEAS model is that the other international partners may be less inclined towards such an ontologically pure model. In particular, some of the more counter-intuitive aspects (such as those mentioned in Appendix A) may be difficult to introduce to data modellers who are steeped in the tradition of relational databases.

3.3 Suggested Way Forward

This study has shown that there is a reasonable mesh between the MODAF concepts (as represented in the early IDEAS model) and the UKDT. Clearly there is a great deal in the UKDT which would not be of everyday use to MODAF architects (e.g. historic estate, medical conditions, etc.), so the use of UKDT in MODAF should be selective. In addition, there will be a requirement to augment the UKDT in the areas of systems and information specifications – the ISSE library may help to fill this gap.

The analysis in this study was a quick attempt to map some of the UKDT terms onto the IDEAS model. A full mapping will require input from CDMA, ontology experts and subject matter experts. It will also require a more formal IDEAS model, and it is crucial that MOD argues the case for a formal ontology at the heart of IDEAS so that it is useful as the MODAF Upper Ontology.

The main issue facing the MODAF project for taxonomy/ontology is one of management. It is clear that the UKDT contains useful terms for MODAF. It is also clear that these must be structured differently if they are to be useful to MODAF architects. If the UKDT covered the entire requirements for MODAF, then this wouldn't be an issue – CDMA could manage the alternative structure in parallel with the existing UKDT structure. However, the UKDT lacks the detail required by architects in the area of systems and information management. This means the MODAF taxonomy will inevitably augment the UKDT, imposing a large amount of additional data management on CDMA. At present, CDMA does not have sufficient resource to manage the MODAF Ontology. However, as the CDMA's Information Layer project starts to mature, it may be the ideal framework under which to manage the Ontology.

In a previous IA document on taxonomy[4], a technical approach based on OWL and Dublin Core (or MOD meta-data standard) was recommended. In short, this approach allowed the UKDT to exist unchanged. The MODAF Ontology would then establish a class structure (in OWL) which re-used UKDT terms by referring to them using Dublin Core (or MOD Meta Data) tags. This approach still appears to be the best way forward as it does not force the MODAF Ontology to be owned by CDMA, however it does raise an issue of configuration management. The UKDT is updated yearly, and it would be important to reflect these changes in the MODAF Ontology – implying an ongoing (though by no means substantial) maintenance task for MODAF. Such an approach would enable the MODAF Ontology to be maintained by another organization – though that organization would have to liaise with CDMA.

Running through the points made above is a clear theme of governance arrangements for the MODAF Ontology - who will own and maintain it ? In addition, there are also questions about day-to-day usage; how will users gain access to it ? how will users be able to extend it for their own projects ? As the MODAF Enablers workstream progresses, the Ontology governance aspects will become significant and will need to be managed as part of the wider MODAF governance. In the long term, it is clear that there is a will within CDMA to integrate the MODAF Ontology into the *Information Layer*. Until that point though, there is no clear governance body for the MODAF Ontology, and this issue needs to be addressed as soon as possible.

4 Conclusions

The IDEAS model has all the right ingredients to be a suitable replacement for the ERM. It is being designed specifically for architectural data, and offers an excellent opportunity for international alignment. The MODAF Meta-Model defines the representation of architectural elements, but the purpose of an upper ontology is to represent the actual things that are represented by an architecture – i.e. the reality. If IDEAS is to be used in this way, it should be founded in formal ontological principles, and this is the main point that the UK part of the IDEAS team will be driving. Since the analysis for this report was carried out, the IDEAS team had a very productive modelling meeting. This has resulted in a v0.2 draft release covering systems and organizations.

With some work, it should be possible to use the UKDT for MODAF (see Section 3.3). This work will require close cooperation with CDMA, with occasional need to request clarification from subject matter experts. The challenge is to mitigate the configuration management issues arising from maintaining two structures of the same classes. One answer to this would be to request CDMA to maintain the pure specialisation hierarchy (Ontology) as well as the current narrowing-terms hierarchy – either in their current taxonomy management tool or as part of their future *Information Layer* project. However, given that the MODAF Ontology will extend the UKDT, CDMA would also have to manage the additional classes (e.g. from the ISSE Library). This may not be practical, so the solution may be for a future MODAF programme to manage the MODAF Ontology in close cooperation with CDMA – i.e. as CDMA issues changes to the UKDT, these can then be incorporated in the MODAF replica. As the MODAF Ontology reaches a level of maturity where the ongoing work is just maintenance, CDMA could incorporate it into the *Information Layer*. The decision on which approach to take will rest on CDMA's ability to take on the extra effort to manage the MODAF additions and structure. The decision may become easier as DG-Info takes on more responsibility for MODAF.

The main recommendation is to proceed with the approach of re-using UKDT elements in the MODAF Ontology by means of Dublin Core (or MOD Meta-Data Standard) references. This allows the two hierarchies to co-exist with a clear indication where the same class exists in each. For the Upper Ontology, the IDEAS model appears to be a good candidate. Although not strictly a formal ontology, it has a clear subtype hierarchy and a reasonable coverage of MODAF (and the coverage will increase with time as IDEAS is further developed). The final decision about using IDEAS cannot be made until a later date – the current version of the model is too immature to judge. However, it is obvious that there are some key elements common to all architectures – system, organization, operational activity, project, etc. A sensible approach would be to start developing the taxonomy from these roots. This could be begun before IDEAS is complete, as there is little or no risk in assuming the existence of those elements. A final recommendation would be for the MOD to suggest a second workstream in IDEAS to develop the upper layers of a taxonomy that could be agreed between the nations. It would be difficult to get agreement on anything more than 2-3 levels deep, but there would be great benefit in even a small international taxonomy.

Appendix A – Background Primer on Taxonomy & Ontology

Classes & Individuals

In discussing taxonomies, it is important to establish a common language for describing the concepts involved. Set theory, category theory and logic provide a technical underpinning for classification. Classes are sets defined by their members, or by some criteria that their members conform to. The membership of a class need not be exclusive - for example, the author of this document is member of the class "human beings", as well as being a member of other classes such as "people born in the 70s", "MOD contractors", etc. One class may be a subset of another – *specialisation*. Taxonomy is the discipline of defining *classes* and how a given class *specialises* into other classes – e.g. "mammals" specialising into "primates", "canines", "felines", etc. Venn diagrams are often the easiest way to communicate relationships between classes.

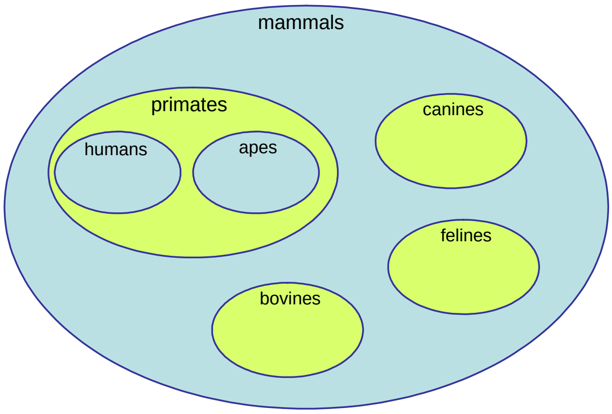

Figure 5 - Venn diagram representing simplistic species taxonomy

(probably would not stand scrutiny from a zoologist)

Individuals are things which belong to classes, but are not classes themselves – for example the author of this document is an individual. Individuals may belong to a number of classes – i.e. classes are identified of which the individuals are members (see Figure 6).

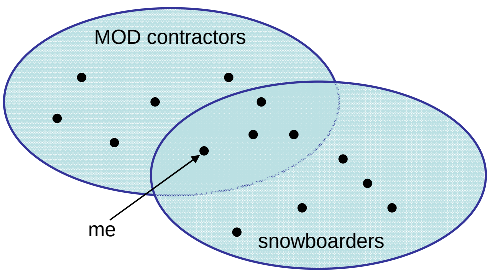

Figure 6 - Individuals belonging to more than one class

However (and this is an important point to remember), the members of classes need not all be individuals – i.e. classes may be members of classes. For example, the class "types of people" has members such as "doctors", "lawyers", "software engineers", etc., which are classes themselves. Most taxonomies do not deal with the case of classes of classes – they tend to define their scope in terms of one class of classes (e.g. animal species). True ontologies, on the other hand, regularly mix levels of abstraction in this way. Traditional information modelling techniques tend to deal with classes and individuals that are members of their classes, and rarely deal with classes of classes explicitly (though there are exceptions – e.g. classes of warship). In addition, taxonomies tend not to deal with individuals, whereas an ontology of any useful merit will contain both classes and individuals.

Ontology & Taxonomy

Ontology is a term that is somewhat fashionable in the IT industry at the moment. However, it is worth dispelling some of the myths that surround ontology. Ontology is a branch of philosophy concerned with analysing the nature of what exists. As a formal philosophical discipline, it has existed at least since the ancient Greeks (Aristotle was probably the first to truly formalise the discipline). In recent years, Ontology has had a more public profile with the emergence of the W3C's semantic web standards – particularly the Web Ontology Language (OWL).

Before explaining what ontology is, it is probably more useful to explain what it isn't – especially to remove some of the "noise" surrounding ontology which has been generated by the IT community. The majority of ontology projects that have been publicised within organizations in recent years are little more than information modelling exercises that simply chose to use OWL as the modelling language. The use of OWL does not automatically create a true ontology. The essence of ontology is the analysis of what exists or could possibly exist – things and the relationships between them. Millennia of effort have gone into refining the discipline. It is somewhat unfortunate that the majority of web ontology projects have overlooked this extensive body of work, because it has resulted in a branch of philosophy where the majority of the difficult problems (that most information modellers "solve" by making arbitrary decisions) have been refined to a level of understanding unrivalled by anything that has emerged from computer science[9]. Many of the tools of ontology (classes, individuals, specialisation, inheritance, class-member relationships, etc.) have already been adopted, though usually in a watered-down form, by formal techniques such as Entity-Relationship and Object-Oriented modelling, and of course, OWL. The way in which these tools are used rarely owes anything to philosophical ontology, and anyone with a formal background in ontology would find little that is familiar in the thousands upon thousands of information models that are churned out using those tools.

An ontology seeks to define what things exist and the ways in which they exist. The key to ontology is taking a "view from nowhere" – i.e. not allowing a particular person's perspective to direct the structure of the ontology. This is where a formal ontology is very different to traditional data modelling and also where it offers a great deal to information science and technology. By allowing systems to manage a more fundamental representation of the world, it enables integration of information from disparate sources, and re-use of common implementation patterns. The key to this is analysis of the (often artificial) concepts which pervade the information we use to establish the true meaning of what is being modelled – i.e. looking at information generated from different perspectives which cover the same subject area. For example, in an information (or data) model it wouldn't be unusual to find an entity (or class) called "materiel requisition record". However, such a compound concept would not easily find a home in an ontology. Instead, the ontology would describe the activity of requesting materiel, relate this activity to the type of materiel requested, and to a document which represents the request. Although this may seem like an unnecessary complication, such an analysis de-couples the description of what is really happening from the perspective-based approach of current database implementations. It also breaks the model down into simple constituent parts (people, activities, documents, etc.) that are common across domains. It is this use of common, immutable concepts that enables integration of information and understanding across disparate parts of an enterprise.

It is important to remember that early software implementations sought to simply automate the existing paper-based processes. This emphasis has remained in software applications to this day, resulting in data models which do not truly represent what is going on, but a paper perspective of the information required to support what is going on[10]. The most common information modelling methodologies favoured by software implementers are based on some form of process analysis. This analysis usually identifies the information flows from one process to another, and uses these flows to drive the development of the information model which will, in turn, drives the structure of the information handled by the software application. This approach is ubiquitous (although the inventors of each methodology would claim novelty, they are all based on essentially the same premise of analysing information flows between processes) and has been very successful at driving out the current requirements of the organization that will use the software. However useful this approach is, it has one fundamental flaw – if there is one thing that is guaranteed to change in an organization, it is process. Organizations of all descriptions constantly re-structure and re-configure themselves to

---

[9] See "Ontology is Overrated" by Clay Shirky - http://shirky.com/writings/ontology_overrated.html

[10] See "The World on Paper". David R. Olson. Cambridge Univ Press. 1994. ISBN: 0-521-44311-3

adapt to new market opportunities and competitive threats. Aside from this obvious flaw, the methodologies rarely prescribe any further analysis of the information to discover the true nature of the information[11]. This inevitably results in the same real-world things being modelled differently in each analysis, resulting in incompatible systems.

Taxonomy is the discipline of classification. For the purposes of MODAF, a taxonomy is a structured classification tree, with general roots that branch into more specialised terms. To ensure that the MODAF taxonomy is useful, it is important to apply an ontologically pure approach in defining the most general level – i.e. the roots of the taxonomy tree. If the concepts defined at this level are not sufficiently "pure", it causes confusion amongst the lower branches of the tree as it becomes apparent that a class could belong in several branches.

[11] There are some recent exceptions to this, and methodologies such as BORO seek to derive ontologies from pre-existing data models and data - http://www.boroprogram.org. The MOD has also recognised the same issue, and sought to mitigate the problems of perspective-based modelling with the development of the CBML language and methodology. This approach allows information models to be first defined from a given perspective, then refined into a more generic model. Although this is a great leap in terms of model quality, the CBML approach is not guaranteed to produce what would be considered a formal ontology.

Appendix B – MODAF Taxonomy Requirements

*The following is re-produced from the IA Document **IA/02/16-ERMcm05** – "MODAF Taxonomy Requirements". Each Key User Requirement is highlighted in red, with a "[KUR]" added after the sentence that describes it. Requirements that simply desirable rather than essential are highlighted in green, with [UR] appended to the sentence. The term MODAF Taxonomy refers to the implementation of the taxonomy, not its content.*

Human & Computer Interpretable

The MODAF Taxonomy shall be defined in a way which can be parsed and comprehended by computer systems [KUR]. This allows the taxonomy to be seamlessly integrated with architectural tools – e.g. a user can select an organization type using the tool of their choice without needing to know that the tool had obtained the allowable types of organization from the *MODAF taxonomy*.

The developers of a taxonomy tend to be domain specialists – e.g. activity modellers for activity types, HR experts for organization & role types, etc. For this reason it is necessary that the format used for the taxonomy is easily understood by non-technical users. It is rare to find a useable, computer-interpretable format that is easy for humans to use. Hence it is inevitable that the human interaction with the taxonomy will be through some form of user interface. *The MODAF taxonomy shall be made available to users and developers via a user interface or editing tool. User access to the taxonomy shall be controlled, with editor and reader privileges [KUR].*

Unique Identification, Availability & Security Classification

In order to be useful to architectural tools and to enable unambiguous data exchange between those tools, there must be a way of uniquely identifying each element in the taxonomy – i.e. when one user chooses a "tank", there must be no confusion with any other elements in the taxonomy. *The MODAF taxonomy shall provide a mechanism for uniquely and unambiguously referring to an element in the taxonomy [KUR]*.

The MODAF taxonomy shall be available on-line to all the parts of the MOD which require access to the taxonomy [KUR]. This may require mirror copies of the taxonomy on secret and restricted networks as well as a "sanitised" copy available on the world wide web for outside users.

It is possible that certain elements or combinations of elements in the taxonomy may be subject to security classification. *The MODAF Taxonomy shall provide a mechanism for identifying which elements are subject to classification, and the level of classification they are subject to. The classification information shall provide the specification for filtering the taxonomy for availability on the various networks – e.g. internet, restricted, secret [KUR]*.

Specialization Hierarchy

The nature of a taxonomy is hierarchy of classifications – e.g. a zoological taxonomy beginning with animal, specializing to mammal, reptile, etc. *The MODAF Taxonomy shall provide a mechanism for specifying which elements in the taxonomy are specializations of other elements [KUR]*.

Distributed Taxonomies

It is unlikely that an organization the size of the MOD can define a taxonomy to the level of detail to which all its stakeholders need. *The MODAF Taxonomy shall provide a set of high-level classifications which are agreed by, and of use to, all MODAF stakeholders [KUR]*. There will remain a requirement for each stakeholder community to manage their own specialist taxonomies, extending the high level taxonomy. *The MODAF Taxonomy shall allow multiple taxonomies to co-exist, all of whose elements shall be directly or indirectly a specialisation of an element in the high level taxonomy [KUR]*. For example, the Army may define their own taxonomy, with an element "brigade" which is a specialisation of an element in the high-level taxonomy "military organization".

In allowing taxonomies to co-exist in this way, it is inevitable that there will be repeated classifications across the organization. Rather than trying to prevent this (which would be fruitless exercise), it is essential that taxonomies be able to “borrow” definitions from other taxonomies. Or, in the case where a repetition is discovered after the fact, it is essential that the duplication can be identified and specified in the taxonomies themselves. The MODAF Taxonomy shall allow one localised taxonomy to use elements of another localised taxonomy. The MODAF Taxonomy shall provide a mechanism for identifying synonymous elements *[KUR]*.

Clearly, the co-existence of multiple taxonomies requires a sensible and selective governance framework.

Instances

Much of the standard data that is used in military organizations is composed of instances as well as classifications. For example, “16 Air Assault Brigade” is an instance of the class “brigade” which is so widely understood in the MOD that is considered standard data. It may be useful for the MODAF Taxonomy to handle instances as well as classifications. At this point (especially if relationships between instances are added) the taxonomy starts to become more an “ontology”. *The MODAF Taxonomy shall allow the representation of instances of classes defined in the MODAF Taxonomy. The MODAF Taxonomy shall allow relationships to be specified between instances provided that those relationships are themselves instances of classes defined in the MODAF Taxonomy [UR].*

To clarify the issue of instances and predicates, it is necessary to provide an example. Suppose that the taxonomy defines the classes “brigade” and “regiment”. Suppose it also defines a class “reports to”. Then it would be possible to create instances called “16 Air Assault Brigade” and “The Parachute Regiment” and link these with a relationship of type “reports to”. In other words, it is important to describe the relationships between types of things – i.e. typical relationships – and the actual things that are of those types.

Governance and Maintenance

Given the distributed nature of the taxonomy, there will be a need for constant checking and control across the organisation. *The MODAF taxonomy shall be under the control of a management committee with responsibility for keeping the taxonomy consistent and resolving any conflicts between taxonomies [KUR]*. In addition, there will be a requirement to manage the content of the taxonomy over time, to ensure it is consistent with current MOD thinking. *The MODAF taxonomy shall be regularly scrutinised with relation to current MOD policy and technology, and corrections / adjustments made where necessary [KUR]*.

A key aspect to making governance work is the way that governance is applied. The goal is a distributed set of taxonomies, each specialising the definitions of others – effectively producing a tree of taxonomies. The governance applied to the root and branch taxonomies would be much more strict than that applied to the leaves. Such an approach allows control of the taxonomy without discouraging individual users to contribute to and use the taxonomy. *The governance framework for the MODAF Taxonomy shall apply varying levels of control on the taxonomy elements, depending up their global vs. local applicability, with the stricter level of control being applied to the global elements [KUR].*

Appendix C – Initial Mapping of UKDT onto the IDEAS Model

This section has been suppressed for two reasons. First, it is likely that the information contained is RESTRICTED. Secondly, it is 400 pages long. The example below is a short, unclassified excerpt from the re-structured taxonomy. The sections numbers are just for convenience and do not represent any formal identification for the terms.

19 System [IDEAS Model]

19.1 Acquisition Management System [UK Defence Taxonomy]

Alternative term: AMS

19.2 Air defence systems [UK Defence Taxonomy]

19.2.1 Air Defence Alerting Device [UK Defence Thesaurus]

Alternative term: ADAD

19.2.2 Ground based air defence systems [UK Defence Thesaurus]

Alternative term: Air defence missiles

19.2.3 Man-portable air defence systems [UK Defence Thesaurus]

Alternative term: MANPADs

19.3 Area defence weapons and weapons systems [UK Defence Taxonomy]

19.4 Avionics systems [UK Defence Taxonomy]

Alternative term: Air data systems Alternative term: Aircraft sensor systems

19.4.1 Air traffic control systems [UK Defence Thesaurus]

Alternative term: Traffic control systems (air)

19.4.1.1 Deployable ATC [UK Defence Thesaurus]

19.4.2 Airborne targeting systems [UK Defence Thesaurus]

Alternative term: Airborne navigation and attack systems Alternative term: Aircraft targeting systems Alternative term: Targeting systems (airborne)

19.4.3 Aircraft display systems [UK Defence Thesaurus]

Alternative term: Display systems (aircraft)

19.4.3.1 Cockpit displays [UK Defence Thesaurus]

19.4.3.2 Head down displays [UK Defence Thesaurus]

19.4.3.3 Head up displays [UK Defence Thesaurus]

19.4.3.4 Helmet mounted displays [UK Defence Thesaurus]

19.4.4 Flight control systems [UK Defence Thesaurus]

19.4.4.1 Airborne collision avoidance systems [UK Defence Thesaurus]

Alternative term: ACAS Alternative term: Collision warning systems

19.4.4.1.1 Fast jet collision warning systems [UK Defence Thesaurus]

19.4.4.1.2 Ground collision warning systems [UK Defence Thesaurus]

Alternative term: GCAS Alternative term: GPWS Alternative term: Ground collision avoidance systems Alternative term: Ground Proximity Warning System

19.4.4.1.3 Traffic Alert and Collision Avoidance System [UK Defence Thesaurus]

Alternative term: TCAS

19.4.4.2 Autopilots [UK Defence Thesaurus]

19.4.5 Flight management systems [UK Defence Thesaurus]

19.4.5.1 Flight data recorders [UK Defence Thesaurus]

Alternative term: Flight recorders

19.4.6 Flight safety information management systems [UK Defence Thesaurus]

19.4.6.1 Health and Usage Monitoring Systems (aircraft) [UK Defence Thesaurus]

Alternative term: Health Usage and Monitoring System Alternative term: HUMS

19.4.7 Landing systems [UK Defence Thesaurus]

Alternative term: Aircraft landing systems

19.4.7.1 Instrument landing systems [UK Defence Thesaurus]

19.5 Business management systems [UK Defence Taxonomy]

Alternative term: BMS

19.5.1 Head Office Business Management System [UK Defence Thesaurus]

19.6 Command and control information systems [UK Defence Taxonomy]

Alternative term: C2 systems Alternative term: C2I systems Alternative term: C3I systems Alternative term: C4I systems Alternative term: CCII equipment capability Alternative term: CCIS Alternative term: Command control and information systems Alternative term: Command control computers communications and information systems Alternative term: Command support information systems Alternative term: NATO command and control systems

19.6.1 Air command and control information systems [UK Defence Taxonomy]

19.6.2 Battlefield information systems applications [UK Defence Taxonomy]

Alternative term: Battlefield reconnaissance surveillance and target acquisition systems Alternative term: Battlefield target identification systems Alternative term: Battlespace management systems Alternative term: BISA

19.6.3 Deployment and movement systems [UK Defence Taxonomy]

19.6.4 Joint command and control information systems [UK Defence Taxonomy]

19.6.5 Land command and control information systems [UK Defence Taxonomy]

19.6.6 Maritime command and control information systems [UK Defence Taxonomy]

19.7 Communications and information systems and CIS equipment [UK Defence Taxonomy]

Alternative term: CIS equipment Alternative term: Communications equipment Alternative term: Information systems equipment