

# Topic Maps for European Administrative Nomenclature

Gabriel Hopmans<sup>1</sup>, Peter-Paul Kruijssen<sup>1</sup>, Leon Oud<sup>2</sup>, Jelte Verhoeff<sup>2</sup>, Marc Wilhelm Küster<sup>3</sup>, and John Clews<sup>4</sup>

<sup>1</sup> Morpheus Software, P. O. Box 240, 6200 AE Maastricht, the Netherlands  
{G.Hopmans, P. Kruijssen}@mssm.nl, <http://www.mssm.nl>,

<sup>2</sup> Conclusion, P.O. Box 85030, 3508 AA Utrecht, the Netherlands,  
{loud, jverhoeff}@conclusion.nl, <http://www.conclusion.nl>,

<sup>3</sup> FH Worms - University of Applied Sciences, Fachbereich Informatik/Telekommunikation  
Erenburgerstraße 19 \* D-67549 Worms, [kuester@Fh-Worms.de](mailto:kuester@Fh-Worms.de),

<sup>4</sup> Keytempo Limited, 8 Avenue Rd, Harrogate, HG2 7PG, United Kingdom,  
[adnomist@uk2.net](mailto:adnomist@uk2.net)

**Abstract.** This work-in-progress report describes the requirements for a "European Administrative Nomenclature" network. It addresses the research topic of interoperability between UN and EU reference classifications, Topic Maps, ebXML Registries and distributed databases. Governmental organisations in Europe will be supported in their administrative terminology with the help of a Seamless Core Model, Published Subjects and TMRAP.

## 1 Introduction

Increasingly, mobile European citizens need to interact with national administrations, and the services that they provide, in several countries (citizen to government, C2G). European administrations need to exchange information between each other (government to government, G2G). Differences between national administrations, their nomenclature, terminology, structures, and centralised information islands approaches make this difficult. These problems will be addressed with the help of Topic Maps within ADNOM [1] (Administrative Nomenclature), a project (or so called Workshop) funded by the European Commission's DG Enterprise through funding available for standardization activities in the framework of eEurope 2005 and is overseen by CEN, one of three European ICT standards organizations (the others being CENELEC and ETSI) from May to February 2006.

The aim of the Workshop ADNOM is to deliver a Workshop agreement, to establish and maintain a network between European government translation units, terminology organizations, etc. with the purpose of developing and disseminating European *Administrative Nomenclature*, built as far as possible on the basis of existing networks and resources.

The prototype ADNOM (Administrative Nomenclature) network, a pan-European semantic resource to increase interoperability in applications and in terminological activities in governmental institutions is planned to be in place by the begin of 2006:

this will provide relevant information to citizens on European and national administrations, in many European languages. ADNOM is a small size project and the goal is to make it a long term network. ADNOM uses Topic Maps and Published Subjects to enable governmental specialists working in eGovernment, terminology and information management to manage the concepts and update the content.

Defining mappings between existing resources, designing cross-lingual resources and guaranteeing reliable distributed knowledge exchange is where a Topic Maps approach will solve several problems. Related work has been done in the SNS project [2] but in this project the advantage of Published Subjects has not been utilized. Published Subjects [3] are a method of establishing semantic identity using URIs in an open, democratic, bottom-up, and distributed process. Standards Norway provides the ADNOM Secretariat. For information on participation contact the Secretary Håvard Hjulstad, who is also chair of ISO TC37 “Terminology and language and content resources” or see [1].

First we will explain the need for ADNOM’s suspected outcome in the working field of Terminology and which resources are already available so that the approaches in the project can be introduced. The requirements that we list here will give the reader an idea why this is a classic application for topic maps. Then we will go deeper into how a topic maps based, cross-border approach to administration can help in establishing terminological interoperability. The last subsection will give insight in how large the impact of ADNOM might be on all Topic Maps activities in Europe.

## 2 ADNOM approaches

Persons working at Governmental organisations in Europe need an administrative nomenclature serving relevant terminology, listing terms with detailed information and explanation. Ideally they should be supported by an interactive system (component c, see later), enabling them to identify terms, make a list of designations, excerpts terms, reduce differences between concepts, harmonise them and process them. ADNOM reuses existing resources and includes in a faceted approach topics like: Governmental function (e.g. Defence, Police, Finance, etc), Jurisdictions (Countries), Organization types (e.g. Parliament, Ministry, Agency). For these, widely used UN [4] and EU reference classifications like Classification of Outlays by Functions of Government, European Nomenclatures (COFOG) [5], Nomenclature of Units for Territorial Statistics (NUTS) [6] are used and it will also build on EUROVOC, AGROVOC and GEMET Topic Maps from the European Parliament Thesaurus available at [7].

ADNOM will provide and integrate the following components with topic maps:

- a) a concise guide to typical government activities in terminology across Europe, part of the standard-type document (a CEN Workshop Agreement).
- b) content on specific government activities in specific countries in this field. (which organisation types, names do we find in a country, which standards, which terminologies do they use, which functions, which languages)
- c) an interactive system for handling topics, terms, codes and content
- d) a meta-terminology for European Administrative Nomenclature

- e) an active registry and repository functionality for the meta-terminology
- f) a long-term network of translators, terminologists, and systems providers

Within ADNOM a lot of principal facets based on existing codes are reused (see [8] and [9]) and it has been recognised which impact facets can have on the power and usability of knowledge resources. The faceted classification as described in Garshol [10] has been followed within ADNOM to cross boundaries between systems and connect terms in a subject-based fashion. The paradigm for the power of facets is described as "Busch's golden law of facets." This states that (in an idealized world) "four facets of 10 nodes each have the same discriminatory power as one taxonomy of 10,000 nodes." (a) faceted navigation helps content owners, as faceted organization enables content owners to streamline their information management processes, and (b) faceted navigation helps users more easily find what they're looking for. The human-computer-interaction (HCI) community correctly suggests that faceted classification doesn't necessarily solve all the problems but with Topic Maps the user can navigate subject based over multiple existing faceted classifications. These advantages will be at the heart of human knowledge organization in ADNOM, and will be reflected in the way that the project develops the semantic infrastructure, knowledge resources, and meets the needs of users.

Topic Maps will allow ADNOM participants to integrate different ontologies, classifications, thesauri, and store terminologies in a language-independent way. A phased expansion of ADNOM to provide information in national languages of all EU countries (including candidate countries) and EFTA countries is planned. Using Published Subjects in combination with a distributed database will allow both humans and computerized systems to make use of the knowledge condensed and integrated in Topic Maps from the multitude of terminology efforts across Europe. For Published Subjects Identifiers (PSIs), ADNOM is using and extending OASIS PSI sets (<http://psi.oasis-open.org>), which currently consist of country codes (ISO 3166) and language codes (ISO 639-1 and ISO 639-2).

The following two components are at the core of ADNOM: (c) an interactive system for handling topics, terms, codes and content and (d) a meta-terminology for European Administrative Nomenclature. This core has been defined as the ADNOM Seamless Knowledge Core (SKC) Model. For the construction of this model a methodology to disambiguate terms and concepts will be followed and this model can be mapped to the Topic Maps Data model quite easily. For the analysis of the methodology the ADNOM project team has followed the steps used in ISO/CD 860.2 (ISO/TC 37 SC 1 N 276): Terminology work - harmonization of concepts and designations. This means to start at the concept level and continue at the designation level, to identify in a feasibility study differences and similarities between concept systems (including contexts), to analyze the context and characteristics of the concept systems, in our case the proposed ADNOM (meta-) nomenclature), containing the material from the then harmonized sources. The comparative analysis of the different concept systems that one has to take into account is really close in how one develops the Topic Maps ontology; one has to identify the relationships between concepts (associations or occurrences); distinguish between the symbol and the thing that refers to it (topics or subjects); think about the depth of structuring; the types of characteristics used to develop the concept system; and the criteria of subdivision used to develop the concepts.

### 3 Terminological interoperability in ADNOM with Topic Maps

When comparing terms with terminological dictionaries one is functionally speaking working knowledge oriented. He uses description of subject fields where one needs to work with terminological entries. These terminological entries are identified by a concept and normally codes or numbers are used for this. For these entries ADNOM defined Published Subjects. It has been recognized that COFOG acts as a high-level structure where the other concept systems (EUROVOC, NACE, NUTS) can be inserted in their appropriate place in the harmonized concept system, forming an overall ADNOM nomenclature. Topic Maps and Published Subjects act as flexible binding points between all these systems. We use a Published Subject Identifier for one term in one system and when one needs to use it on another system we can use TMRAP [14] or merge Topic Maps.

Only adopting a system like NUTS is not sufficient enough within ADNOM since it will break in what we call the Seamless Knowledge principle. This principle is that one can make a statement and use an indicator without being forced to stick to one hierarchy or one subject within too many boundaries, the user must be able to choose or publish flexible and bottom up. For example a user working with NUTS will quickly face the problem with different systems of internal boundaries where, e.g., say Croatia has no regional division. This example applies cross domains in much of eGovernment. ADNOM will ensure that it is flexible enough to accommodate these differences in the concept systems within the administrative domain, and one important step towards achieving this goal – namely to allow for cultural differences rather than making a country fit a model – goes via the use of topic maps. With the ADNOM approach using topic maps will fit the concepts and terms, rather than concepts and terms being forced to fit the model, the past problem within administration.

An Administration Module that is running on the Ontopia Knowledge Suite (OKS) will maintain the ADNOM core topic maps. ADNOM terminologists will firm up in additional concepts in these existing European codes, and deal with the maintenance of these, both during the project, and in its long-term phase after the initial phase.

Figure 1 shows an example how a user within ADNOM can connect several terms with fragments of facets for functions, jurisdictions, organs and the role of Published Subjects and Topic Maps. The example will be motivated next, just note the hierarchies for the faceted classifications in the upper part. For most hierarchical relations ADNOM will use the PSI's available at [11]. Multiple hierarchies used in different systems and maintained between several organisations can be navigated seamlessly. For the organisations it becomes possible to integrate their Web-portals since their content is partly and becomes completely subject based. The concept "Parliament" appears here in two hierarchies: in the Administrative functions hierarchy and in the Government hierarchy (in the lower part of the figure) defined by ISO 860 CD2 Terminology work - Harmonization of concepts and terms. ADNOM will also implement some parts of the meta-model from the Terminology Markup Framework of ISO 16642:2003 to express structural organisation of meta-model, to specify data categories and how they relate to meta-model and to indicate which vocabularies are used.

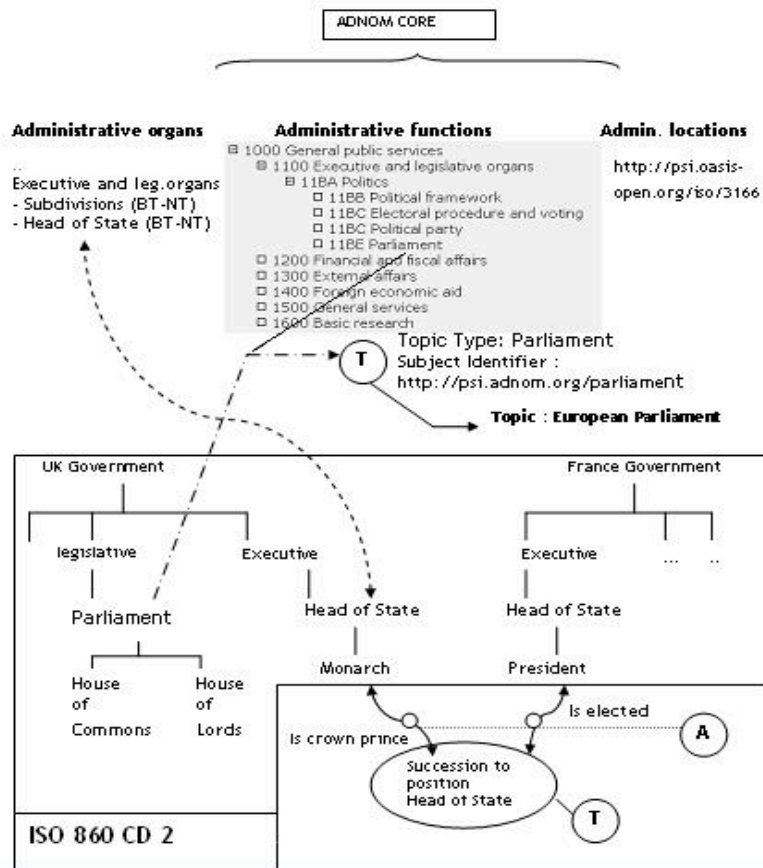


Figure 1: ADNOM core Topic Maps and Published Subject to interrelate hierarchies

The hierarchies for UK and France are connected to the ADNOM levels with Published Subjects. The lower part illustrates how one can make assertions between hierarchies. In this example we wish to designate or represent a term "Succeed" to assert the succession of the position "Head of State" and how to become one. The ADNOM functionalities will enable organisations to publish their terminology work on-demand; they can make statements around subjects with their point of view, terms that organisations need to share can be published with functionalities that have similarities with RSS news feeds like TMRAP [14]. In the ADNOM nomenclature we have for instance the identifiers <http://psi.adnom.org/code/a11BA> for the term "Politics" and we convert this same identifier when doing upconversion of the EUROVOC thesaurus where "Politics" appears as well to connect the systems.

Different assertions about concepts and terms can be made by several terminologists and can be selected from the PSI sets in the distributed repository. Reaching group consensus on the published concepts is then for next stages in which increasingly more and more stable and trustworthy PSI sets will be developed for European Administrative Nomenclature.

## 4. Conclusion and Further Work

Using Topic Maps and a faceted approach for existing classifications, thesauri, and vocabularies enables the users in the ADNOM project to identify their own terms and to harmonise, list and process them for their own activities. By reusing reliable existing resources like those of the COFOG and NUTS nomenclatures, ADNOM already can list terms in 23 languages (see part of this list at [8]). For registry and repository functionalities for European Administrative Nomenclature, the effectiveness of Topic Maps will be combined with other approaches such as the ebXML registry service. The project results with ebXML will be compared with the Topic Maps Remote Access Protocol (TMRAP). Organizations willing to exchange knowledge and that are using Published Subjects will be able to get information from remote repositories of those other organizations automatically. First efforts in the ADNOM workshop need to show how TMRAP/ebXML services will fuse in the architecture. These technologies will probably prove to be orthogonal approaches.

In ADNOM Terminologists will be able to make concept-oriented equivalences like presented in [13]. Nowadays there are over 50 nationally recognised languages in the EU with historical, cultural and legal differences among the various states and regions. CEN emphasises the importance of a long-term project undertaking to link government nomenclatures and terminologies throughout the EU. Input into ISO, W3C and OASIS standards are also planned. ADNOM's influence may also be felt in eGovernment circles during the EU presidencies of the UK (2005) and Austria (2006), and long-term development of ADNOM after that is also planned.

## References

1. CEN/ISSS ADNOM Workshop <http://www.cenorm.be/iss/ADNOM/>
2. Bandholtz, T. : Final Report Implementation of a Semantic Network Service (SNS) in the context of the German Environmental Information Network; 2003, <http://www.semantic-network.de/sns-summary-2003-06-06.pdf>
3. OASIS Published Subjects Technical Committee Recommendation <http://www.oasis-open.org/committees/download.php/3050/pubsub-pt1-1.02-cs.pdf>
4. United Nations Nomenclatures <http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=4>
5. European Nomenclatures, <http://europa.eu.int/comm/eurostat/ramon/>
6. NUTS, [http://europa.eu.int/comm/eurostat/ramon/nuts/home\\_regions\\_en.html](http://europa.eu.int/comm/eurostat/ramon/nuts/home_regions_en.html)
7. European Parliament's "Interoperability Forum" multi-lingual thesauri, [http://www1.europarl.eu.int/forum/interop/dispatch.cgi/sw\\_km\\_lib](http://www1.europarl.eu.int/forum/interop/dispatch.cgi/sw_km_lib)
8. ADNOM COFOG coded nomenclature, codes edited by John Clews and information about Terminology and Nomenclature conference available at <http://mssm.nl/materials/adnom/>
9. ADNOM Seamless Knowledge Core Model, <http://www.adnom.org/adnom-sk/>
10. Garshol, L.M. : Metadata? Thesauri? Taxonomies? TopicMaps! Making Sense of it all; 2004, Paper available at: <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>
11. Ahmed, K. : Topic Map Design Patterns For Information Architecture, XML 2003; 2003, Paper available at <http://www.techquila.com/tmsinia3.html>
12. Wright, S.E. : The basic principles of Terminology Management, Open Forum 2005 on Metadata registries; 2005, Paper available at <http://www.berlinopenforum.de/>
13. Garshol, L. M. : TMRAP - Topic Maps Remote Access Protocol; 2006, In Proceedings of TMRA'05, LNCS, Springer