Title: A7.9-D2 Profiles for border security (Interim version)
Title: A7.9-D2 Profiles for border security (Interim version)
Author(s)/Organisation(s):
Marian de Vries (TUD), Dániel Kristóf (FOMI), Gabriela Szomolaiová (INGR–CZ), Jerzy Wiśniewski (INGR–CZ), Sisi Zlatanova (TUD)
Working Group:
WP7
References:
771-hs_border_security_-_scenario_specification-ingr-001-new 920-hs_border_security_data_specifications-fomi-001-new.xls

Short Description:
This document contains the interim version of the data specification of the Border Security Scenario. It is based on the WP9 system specification and the further discussions about the best fitting conceptual data model for the described use cases.
Keywords:
Border Security, information needs, conceptual data model, harmonisation issues

History:

Version	Author(s)	Status	Comment
000	Marian de Vries	new	Based on first inventory of information needs and harmonisation issues of Scenario
001	Marian de Vries	rfc	Data model for intrusion detection, event response and reporting
002	Jerzy Wiśniewski, Gabriela Szomolaiová, Sisi Zlatanova, Dániel Kristóf, Marian de Vries	update	Update of data model, based on Delft meeting and later discussions
003	Dániel Kristóf	update	Added descriptions of available data sources for Hungary
004	Marian de Vries	final	Added data model for topographic and other data, final edits

Table of contents

1 Purpose and scope

This document contains the interim version of the data specification of the Border Security Scenario. It is based on the WP9 system specification and the further discussions about the best fitting conceptual data model for the described use cases. The document also discussed some of the open issues in the Scenario, especially related to the availability of data for the region that was originally chosen as test area (borders of Hungary and Slowakia).

In this interim version the focus is still on the particular use cases of the Border Security Scenario. This was needed to get a better view on information needs, and data harmonisation issues in case of cross-border cooperation between countries. Because of the nature of the application domain (security, safety) and the type of agencies that are normally involved (military, police, government) it was not possible yet to find re-usable data models.

There are a number of European and international agencies that deal with border security, but their publications do not contain detailed enough information to serve as example or as some kind of standard for the Scenario. This lack of published information about border security systems and data flows is ofcourse due to the specific nature of the border security application domain. There has been contact however with a small group of potential users at the ministries of the interior in Hungary and Slowakia, to get input about user needs and data requirements.

During the development of the conceptual data model it became clear that not all data sets needed for the use cases would be available, particularly sensor data sets that contain (raw or filtered) data about border intrusion and other events along the border. It is therefore not sure at the moment whether all parts of the data model presented in chapter 3 can be tested with real data. In that case the option must be discussed whether the sensor data streams can be simulated, so that a test can still be carried out.

A second remark is related to the topographic and other reference layers that will be used in the Scenario. These kind of spatial objects are for clarity reasons moved to a separate package in the conceptual data model, to have a distinction between these (static) data and the dynamic operational data described in the other package.

2 Information analysis

2.1 Use cases of the Scenario

The overal Scenario application functionality is: detection of incidents along the border of Slovakia and Hungary (border intrusion, of all kinds: illegal entry, smuggle, security endangering activities).

The following use cases have been specified in the WP9 Report:

HS01-01: Topographic Data Management

HS01-02: Human Intrusion

HS01-03: High Level Management

HS01-04: Strategic analysis

HS01-05: Strategic Planning

The first use case prepares for the other 4 uses cases: in this topographic data management use case the static datasets of the countries are harmonised with each other, in preparation of the other 4 use cases.

Use case 2 and 3 are the operational use cases (border crossing incident response), and the last two (4 and 5) can be caracterized as ‚analysis, reporting and policy preparation' tasks.

2.2 Spatial information used in the Scenario

The table below gives an overview of the information items relevant in the Scenario, and the purpose for which the specific (spatial) information is used (input) or produced (output).

Information item	**Purpose**
Location, names and types of **transport network**	Getting information on **road, street, railroad network**: ▪ Basic data set for spatial analysis and strategic planning ▪ Basic data set for defining pre-configured procedures
Location, types and names of **rivers**	Getting information on **hydrography**: ▪ Basic data set for spatial analysis and strategic planning
Land cover / Aerial images	Getting information on **open and built-up areas**: ▪ Providing information on types of land cover (forest, settlements, water, farm land …) ▪ Basic data set needed for orientation purposes
DTM	Getting information on **elevation ratio**: ▪ Detecting possible ways of intrusion and breakaway ▪ Basic data set for spatial analysis and planning
Sensor data on	Getting in situ information of the **crisis situation**:

the border	▪ Data provided by digital sensors or human about the crisis situation (phone calls, images, videos etc.)
Internal secured data	Getting information on **patrol units**: ▪ In situ data - position, equipment of the patrol units, number of unit members etc. ▪ Information on transport equipment
Pre-defined procedures according to specific constraints of the event	Getting information on **process** to handle the crisis situation: ▪ Processes are pre-defined in the system according to specific constrains ▪ On basis of the procedures C&C centers are able to handle the crisis situation quickly and effectively
Output: Maps and tables in reports using harmonised terminology and graphical representation	To have comparable statistics and maps about border intrusion and other security related incidents

3 Common data model

Based on the use case descriptions and information analysis phase the following sketch of the common (conceptual) data model was made, consisting of two packages, one for the operational data (Figure 1) and a separate one for the topographic and other reference data (Figure 2).

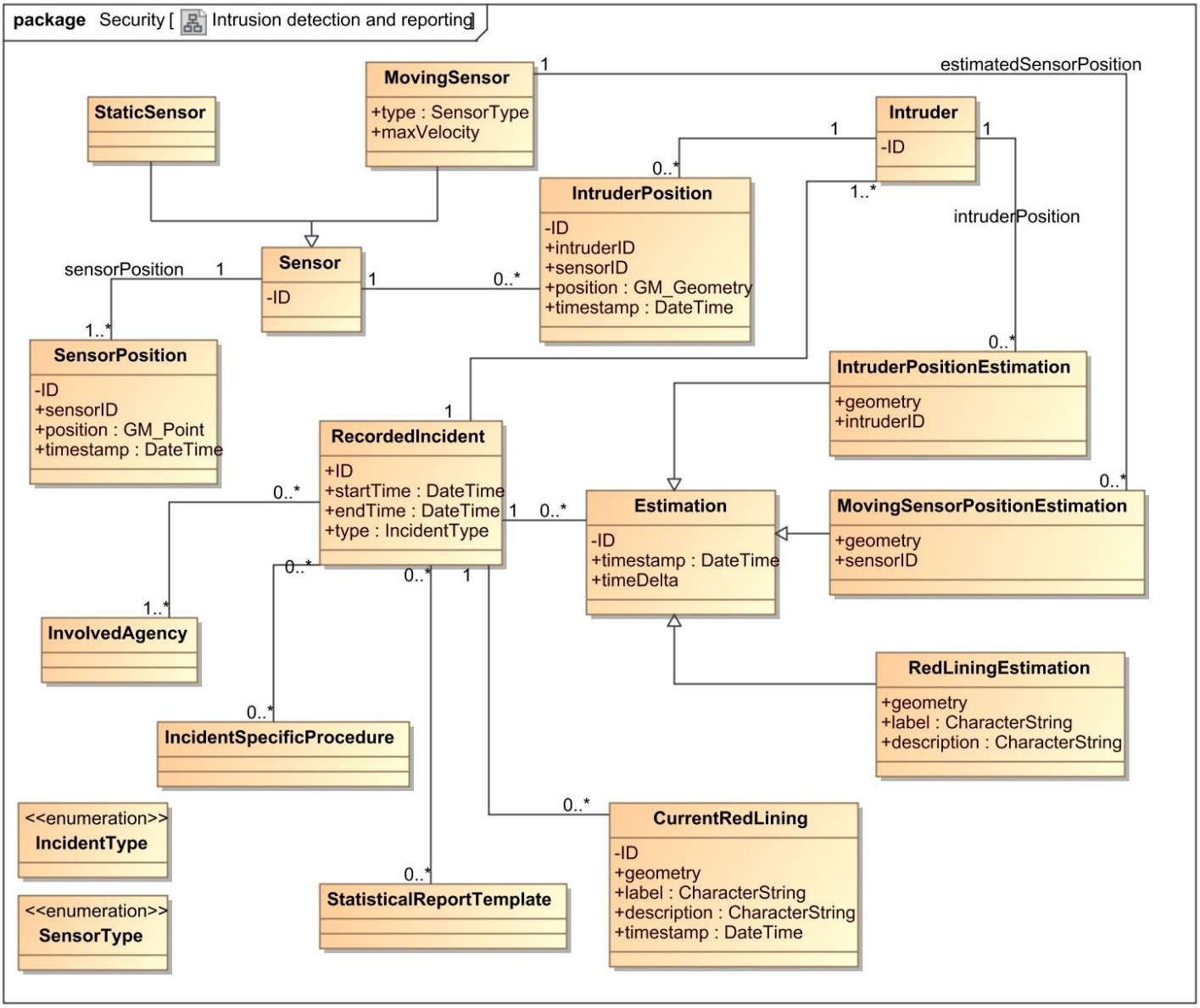

Figure 1 Data model, intrusion detection and reporting package

The ‚intrusion detection and reporting' package specifies the dynamic data: the captured events, and the reports and continuous updating of the location of the ‚actors' (intruder, patrol unit) (red lining). It also contains a class as placeholder for a document system or database with procedures and specific tasks for each type of incident (i.e. the IncidentSpecificProcedure class).

This version of the model does not contain all details yet, but gives the main structure and content of the operational data.

The second package describes the topographic and other reference data, see Figure 2.

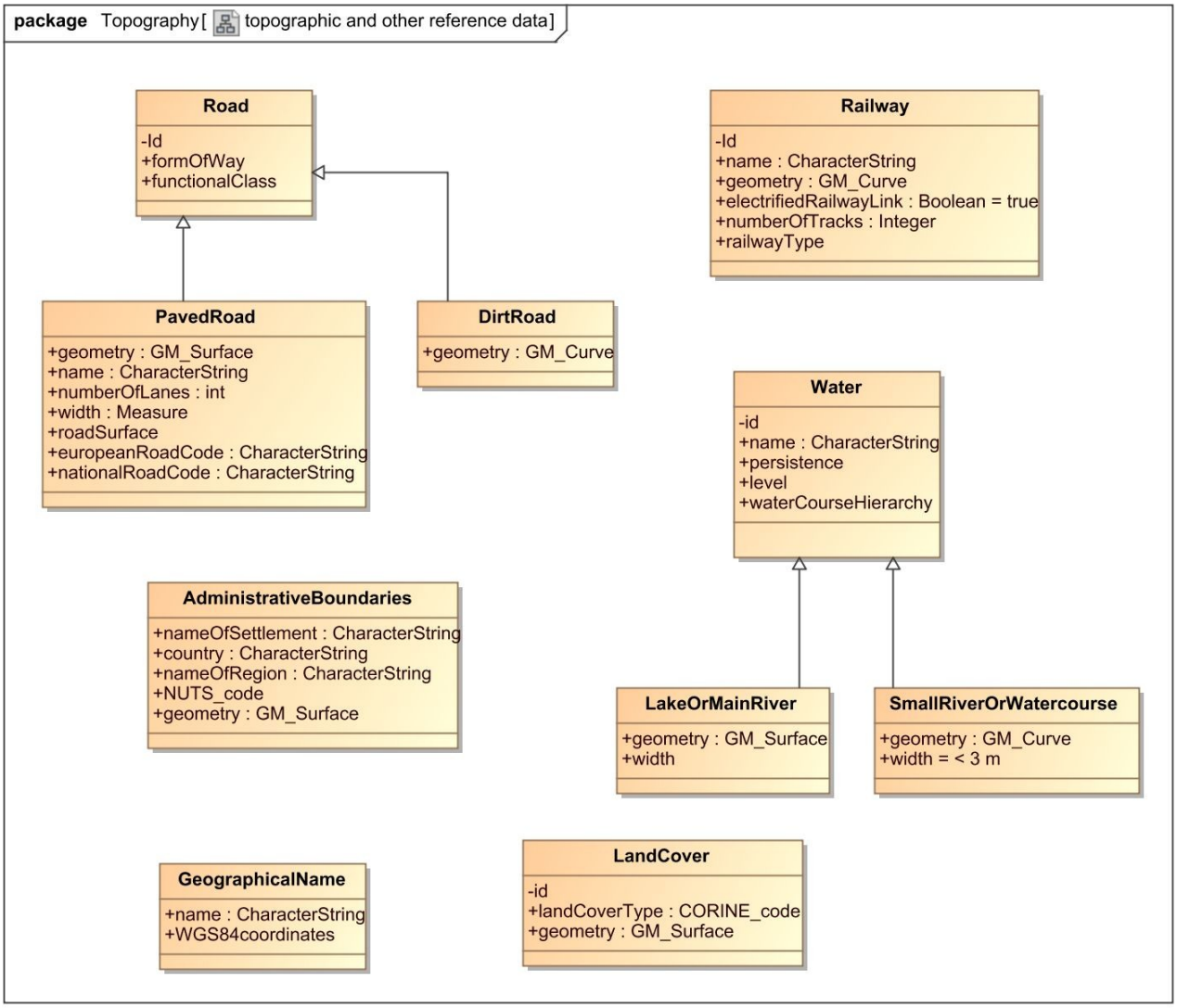

Figure 2 Data model, topographic and other reference data

This second package specifies the static information used in the Scenario, as background layer for orientation and navigation, for spatial analysis of physical border conditions, and for creating maps and statistical overviews in the reporting phase.

For this ‚topography and other' package a number of INSPIRE Annex I theme data models are relevant, that is: Hydrography, Transport Networks, Administrative Units and Geographical Names. A number of attributes of the classes in the package are re-used from these Annex I data specifications, but complete INSPIRE spatial object types have not been re-used in their entirety, because of lack of corresponding data items in the available data sources (see chapter 4).

3.1 Rationale

The 'intrusion detection and reporting' package of the conceptual data model (Figure 1) could not be verified (yet?) against existing standards related to border security, for the simple reason that we did not have access to this kind of government and military information (also see chapter 6). However, what will be done in the final version of the data model is to compare the event monitoring and reporting part to possibly relevant data models in the field of crisis response in general.

A relevant initiative that could be the source of additional ideas for the conceptual data model is the Information Systems for Crisis Response and Management (ISCRAM) discussion and research community, see http://www.iscram.org

Secondly, for the sensor part, the OGC standards and test beds in the field of sensor data management are of importance, such as the SensorML and SOS implementation specifications (see references in chapter 7). Special attention will be paid also to the architecture and protocols proposed and tested in the European projects Orchestra and SANY.

In that way border security becomes a special case of crisis response, and of real-time, sensor data management respectively. For the crisis response part the emphasis is on data harmonisation (maps, terminology, portrayal) and procedure harmonisation plus digital storage of and access to the tailored procedures and tasks per type of event and level of urgency. For the sensor part the focus is then on handling dynamic data, and how to combine data streams from different sensors controlled by different agencies in real-time.

4 Data sources

This is the list of data sources needed in the Scenario of which the availability is certain, and where access rights do not pose a problem. For Hungary the situation is better than for Slowakia.

For a complete overview, with more details per data source, see document: 920-hs_border_security_data_specifications-fomi-001-new

Information item	**Data sets**
Location, names and types of **transport network**	Paved roads: topo_10k_polygon Dirt roads: topo_10k_line Railways: topo_10k_line (all vector)
Location, types and names of **rivers**	Lakes and main rivers: topo_10k_polygon Small rivers and watercourses: topo_10k_line (all vector)
Land cover	CLC50_HU (vector)
Topographic map	Topo_10k_ras_HU (raster)
Aerial images	Orto50cm_2005_HU (orthophotos)
DTM, elevation	DDM5_HU (raster)

Administrative boundaries	For Hungary: admin_bdrs_HU (vector) For Slowakia there is a WMS service with administrative boundaries at: http://www.geonet.sk/main_en.htm
Geographical names	FNT (Gazetteer of Hungary) (MS Access database)
Sensor data on the border	No ‚real' data available, possibility is to create simulated sensor data
Internal secured data	Same as above
Pre-defined procedures according to specific constraints of the event	Textual format or queryable database, to decide. How to implement this can be adopted from examples from other event/crisis response application domains.

5 Harmonisation issues

This chapter is based on the initial inventory made for the WP7 and WP9 reports. Some points were updated or further clarified following the later discussions and the list of requirements specified by the Scenario as input to WP5 and WP8.

5.1 Data format

The commonly used data formats can already be read directly by GIS client systems, others need to be converted first.

This topic should be solved by harmonisation in the HUMBOLDT framework, where an expert user would be able to choose what the output data format is.

Harmonisation of the data format can be achieved via standardised web services (WMS, WFS).

5.2 Spatial and temporal reference systems

5.2.1 Spatial reference system

The spatial reference systems differ between Slovakia and Hungary:

Slovakia has: S-JTSK (Systém – Jednotné Trigonometrické Sítě Katastrální = System – Uniform Trigonometrical Cadastral Network), which is EPSG: 2065

For the Slovakian DTM dataset the vertical datum used is: BpV (Balt po vyrovnaní = Baltic Vertical Datum-after Adjustment).

Hungary has: HD72 / EOV (Egységes Országos vetület = Uniform National Projection), which is EPSG:23700.

Current solution:

The differences in spatial reference systems are partly solved in the GIS client, which manage the coordinate transformation process internally (if the coordinate systems are known, and defined in the system).

Approach in the Scenario:

Relevant for the Scenario are the following SRS systems: EPSG 102067, 2493, 2494, 28403, 28404, 32633, 32634 and 4326.

Coordinate transformation between these systems should be supported by the software used in the Scenario, either in the GIS clients, or in the WMS/WFS services, or with a separate transformation service.

5.3 Conceptual data model

Yes, the conceptual data models for the basic geographic reference information differ between Slovakia and Hungary. Adding more countries will increase the differences in data models.

Current solution:

The differences in data models are solved manually or with semi-automated tools for model harmonisation. These solutions can only be used by expert users, which is a disadvantage.

5.3.1 Geometry types

The ‘intrusion detection and reporting’ package holds the sensor data and has to deal with moving objects. A flexible way to construct tracks from point measurements is needed. But otherwise there are no complex geometry issues, because only the point, polygon, and line geometry types will be used.

5.3.2 Importance of time / Dynamic data

For the sensor data this is of course very important. It is necessary to harmonise the way time and date are specified, and how history of the (filtered) sensor data is kept. There are examples in international standards, which will be followed in the Scenario where possible.

5.4 Classification

Yes, the classification scheme issue is relevant, for example for the distinction between crisis levels (severity of the incident) because of the cross-border communication. And also for the static data, such as the land cover classification.

Harmonisation of classifications is partly dealt with through the specifiation of the common data model, which will include common code lists, and partly by constructing additional terminology listings or thesauri.

5.5 Terminology

Yes, the differences in used terminology/vocabulary are a relevant issue for the cross-border information exchange and cooperation in the field in responding to indicents along the border.

Approach in HUMBOLDT:

Implementing catalogues dealing with used terminology from all involved countries.

Once this is done, also the legends and symbols of the maps and the terminology used in the reports can be harmonised, using the correct vocabulary and graphical representation, also see 5.8 Portrayal.

5.6 Metadata

Metadata in HUMBOLDT should be in compliance with INSPIRE and ISO. For the Scenario it is in addition important that the metadata is shown to end-users in a "human readable" format.

5.7 Scale, aggregation

Detailed reference maps are needed to accurately position the border security events. Therefore 1:10,000 maps (vector and raster) are a minimum condition, having a larger scale would even be better.

It would be good to have the same map scale on both sides of the border. If this is not the case, automated map generalisation, could help. However, for the operational phase, automated generalisation of (vector) maps does not have a high priority. Also aerial photos or satellite images can be used for navigation and orientation.

For the reporting phase, thematic (and geographic) aggregation will be useful: number of incidents can be collected per small region, but also at the provincial level, or for the country as a whole. This can however be handled by the normal existing GIS tools.

5.8 Portrayal

Every user/organisation uses its own style and legend for visualisation. Currently the differences are solved manually by defining a certain portrayal style.

In HUMBOLDT, in the Border Security Scenario, users solving crisis situations should have the ability to choose from pre-defined portrayal styles that are then applied to the datasets of the different countries that have the same thematic content. This interoperability will make the communication easier, e.g. between different local Command centres.

The common portrayal style should also include the map legends. These legends should anyhow be clear for non-expert users. Language translation / localisation is an option for the legends (see 5.10).

5.9 Processing functions

Not in the strict sense of algorithms and parameters, but in the sense of having the same thresholds and warning levels. And secondly in having agreed-upon procedures how to respond to certain incidents: the crisis response organisational aspect. Specifying these common procedures is out of scope for the Scenario, but a database or document system with procedures and tasks specific to certain situations is an element in the common data model (see chapter 3).

5.10 Multi-linguality

Yes, this is an issue, because the Border Security agencies in Slovakia and Hungary use their systems in different languages. Translation of important terms in incident response is needed. Implementing international catalogues for the used terminology (dictionaries) could help. But in an operational situation (in the field, or in a Comand centre) it is too slow when someone has to first manually look up the relevant terms in the agreed-upon common language. Automated language translation is then an option: the application software should be able to execute language transformation of the used terminology in different languages to the target language.

A reverse strategy is that translation is sometimes warranted the other way around: from common language (English?) to the local language, which could be of benefit in case of non-expert users of for example land use maps.

5.11 Spatial, temporal and thematic consistency of data

5.11.1 Spatial consistency

This is an important data harmonisation aspect for Border Security. Data from adjoining regions, with potentially overlapping areas, must form a consistent layer (data set), to avoid misunderstanding. Having software and best practices for edge matching is therefore very relevant.

5.12 Priorities

Both levels of harmonisation (the general, data model and complete data set, level, and the individual spatial object instance level) are important in this Scenario.

High priority is given to: common terminology and portrayal, language translation (multi-linguality aspect), and to spatial and thematic consistency (seamless and consistent map layers using edge matching techniques).

All of these are important preconditions for fast and consistent decision making in operational situations. For the reporting phase again terminology and portrayal is important, and in addition having the possibility of thematic and geographic aggregation into statistical results at different aggregation levels.

6 Open issues

In this chapter some general issues are collected that must be discussed and dealt with in WP7 and WP9 for this Scenario.

6.1 Internal data

In the use case description there is also data/information mentioned as ‚internal secured'.

The question is whether this data needs to be harmonised (and can be harmonised). Can it remain as it is (because it is not exchanged between the border security organisations?). Or is it necessary to also look at this internal data because of the strategic analysis and reporting use cases?

6.2 Access to needed datasets

For the availability of data sets a difference must be made between the data about border incidents, (both dynamic sensor data and historical data in reports) on the one hand, and geo-data such as topographic reference layers or information about administrative regions and responsible agencies.

This is due to the fact that (sensor) data about border incidents are owned and used by the government for military and public safety purposes, and thereby restricted, so that this data cannot be used by the Scenario. An idea is to simulate the sensor data streams about border intrusion events. This could help in realising the full range of the use cases. For this simulation also a series of GeoRSS messages can be used, that are fed to the demonstrator application in real-time. If this is not possible however for practial reasons, the emphasis must shift to the reporting use cases.

As far as the availability of the static reference data is concerned, there is difference still between the Hungarian and Slowakian situation. At the moment there is enough topographic and other reference data available for Hungary. For Slowakia it proves to be more difficult to discover and obtain the needed data sets.

7 Related projects and initiatives

DHS (US Department of Homeland Security) Geospatial Data Model, http://www.fgdc.gov/fgdc-news/geo-data-model-v25

EuroBoundaryMap, http://www.eurogeographics.org/products-and-services/euroboundarymap

European Border Surveillance System (EUROSUR)

FRONTEX, http://www.frontex.europa.eu/

Havlik D., Bleier T., Schimak G. Sharing Sensor Data with SensorSA and Cascading Sensor Observation Service. *Sensors*. 2009; 9(7):5493-5502

ISCRAM (Information Systems for Crisis Response and Management), http://www.iscram.org/

LIMES (Land and Sea Monitoring for Environment and Security), http://www.fp6-limes.eu/

OGC (2007a). OpenGIS® Sensor Model Language (SensorML) Implementation Specification, Version: 1.0.0. OGC® 07-000

OGC (2007b). Sensor Observation Service, Version: 1.0. OGC 06-009r6

OGC (2007c). OpenGIS® Transducer Markup Language (TML) Implementation Specification, Version: 1.0.0. OGC 06-010r6

OGC Sensor Web Enablement WG, http://www.opengeospatial.org/projects/groups/sensorweb

SANY Sensor Anywhere - IST FP6 Integrated Project, http://sany-ip.eu/about_sany

SeBoCom (Secure and Interoperable Border Communications)

Usländer, T., RM-OA - Reference Model for the ORCHESTRA Architecture Version 2 (Rev. 2.1). OGC Best Practices Document 07-097