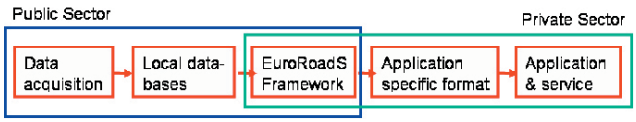
Data Processing

Introduction

The key output of the EuroRoadS project is the specification framework which enables the provision and exchange of harmonised, interoperable and quality assured road data. It provides a vital link between potential (public sector) data suppliers and (private sector) data users.



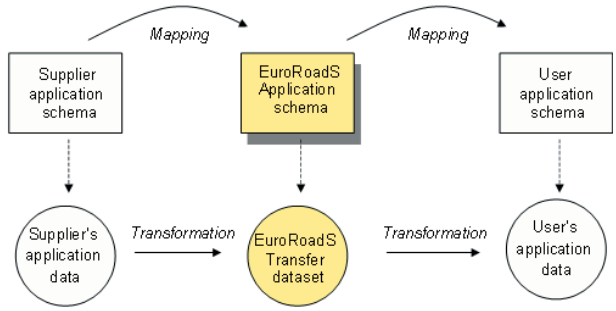
The EuroRoadS data chain

The specification framework needs to be implemented and tested in order to demonstrate and validate its usefulness. This is subject of the demonstration work in the project which establishes an exemplary information chain for speed limit data from data acquisition to a final end user application.

This leaflet summarises the demonstration activities related to data processing for data supply by different public authority providers and data use by the application provider. Furthermore, the demonstration activities are described in the leaflet 'demonstration overview' and leaflets describing 'data capture and maintenance', 'metadata server' and 'client application SpeedAdvice'.

Objectives

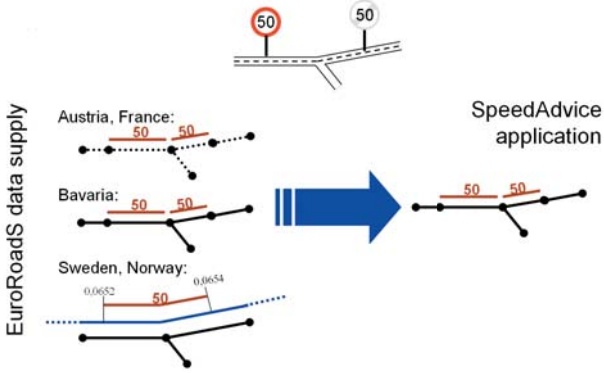
In general, the demonstration activities shall validate the EuroRoadS specification framework. This means specifically that experiences with transfer and use of similar data from different (public authority) sources shall be generated and documented at the example of a show case for a concrete data chain. Speed limit data have been chosen for this show case and shall be tested in a mobile end-user application providing speed limit information while driving ('SpeedAdvice').



Logical mapping, transformation and physical transfer of data

Principle Steps in Data Processing

The data processing starts on the supplier's side. The supplier's own local information model (or application schema) needs to be mapped logically to the EuroRoadS information model. This can include the definition of certain transformation steps if no data object/attribute is available, which corresponds directly to those required by the EuroRoadS data model. The local data needs then to be transformed and exported to the EuroRoads exchange format (XML).



Speed limits referencing and network description by EuroRoadS data suppliers in the demonstration.

On the data user's side following steps can be discerned:

- Again a logical mapping for the EuroRoadS model to (usually proprietary) object schema has to be done,
- The EuroRoadS exchange data set (XML) has to be imported into an internal data base (and schema) on which all further operations are executed,
- The supplier has to do some data merge and integration operations, this is the core part elaborated further below,
- Finally, a conversion into the application specific format has to be done. In case of the SpeedAdvice application this is a binary format which is optimised for mobile devices with regard to performance and data volume.

Every EuroRoadS supplier has to provide at least the mandatory data set comprising a road network description and the speed limit content.

The EuroRoadS information model allows different ways to represent these data. Key differences between suppliers include (see figure below)

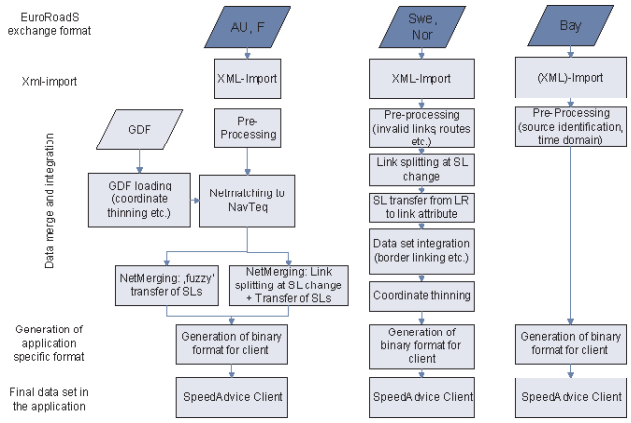
- different network representations (geometry only, geometry + topology) and
- different ways to attach specific content such as speed limits (direct link attribute or linear referencing to route)

Hence , the data merge and integration operations of the speed limit data by the data user (application provider) differs for each supplier. These processing steps have to convert the incoming data and combine them if needed with external data to meet the final application requirements. For the demonstration the final application SpeedAdvice requires a complete, routable network description (geometry + topology) with speed limits attached as attributes to each individual link.

Two scenarios for data use: 'Selective' versus 'Exhaustive'

There are two basic scenarios for using the network as well as the 'attached' data objects provided through EuroRoadS:

- The supplied EuroRoadS data set is exhaustive for the application, i.e. no further data need to be integrated for running in the application. This requires sufficiently harmonised data sets on the supply side and sufficient detail of the network description depending on the application. In the demonstration example the minimal network description required by EuroRoadS (geometry only) would usually not suffice to meet the routing re-quirements of the application.
- Only the additional objects are used selectively from the data set provided. The network supplied only serves to provide the referencing bases. In this case, only a very limited network description is needed. In case of different national suppliers only the speed limit data need to be supplied in a harmonised way.



Data processing steps for the data user

The exhaustive data use is trialled for the Swedish, Norwegian and Bavarian test areas, the selective scenario is trialled for the test areas in France and Austria.

Data merge and integration

In the following, the main processing steps are summarised for the different data sets processed.

France/Austria

France and Austria supply geometrical road informa-tion (no explicit topology) only. The speed limits are handled as road link attribute, i.e. for each road link one speed limit information is attached. Since the network description is not sufficient for the applica-tion, the speed limits are transferred to a appropriate road network which is derived from commercial map data (GDF). The transfer of references is made pos-sible by using net-matching techniques, which allow to identify corresponding links in the supply network as well as the target network. In a separate step, the speed limits are transferred to the target network ('net-merging'). This can be done in a simple way, where the best corresponding link (approximate matching) is used or in a more sophisticated way where links are split in the target network to create ideal matching pairs between supplier and target network links.

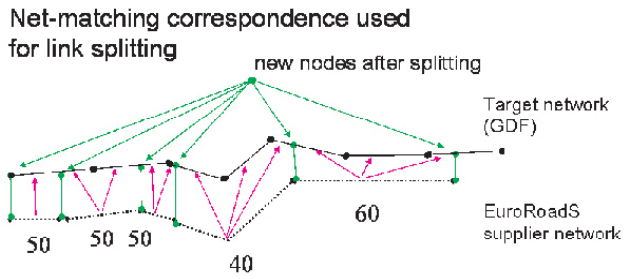
Bavaria

The Bavarian data set is based on a GDF derived data model and content, which makes it very close to the required data input for the speed advice application. The data set (geometry, topology of the network + speed limit data as road link attribute) is therefore exhaustive for the application and only little data processing is needed before the conversion into the final data format can be done. Since the data collection was done in the context of EuroRoadS the detail in speed limit data are relatively high (e.g. distinction between rule based and road-sign speed limits, time dependency, weather dependency etc.). Also, in contrast to all other providers, a link can 'carry' more than one speed limit (e.g. the rule based and the road-sign based). The handling of different types of speed limits per link is done by the final application.

Sweden/Norway

The Swedish and the Norwegian data set also are exhaustive for the application and do therefore not require any external data set to be merged. The speed limit data are not directly referenced to the link, rather they refer to distances ('offsets') along a route, which itself is typically a chain of links ('linear refer-encing' or 'segment expression' in EuroRoadS jargon). Hence, speed limits typically do not start or end at a node. Therefore the existing links need to be split at the appropriate positions and additional nodes need to be inserted ('Link splitting at speed limit change'). Only then, the transfer of speed limit information from the linear reference to the route to a road link attribute can be done.

In a second step, the two data sets need to be inte-grated at the common border ('border linking'), which means that a topological (but not necessarily geometrical) connectivity between the links on either side of the border need to be established. This can be done in different ways by the data user. In demonstration, the data user's internal format and application use the concept of a 'teleporter' object (border node layer), which holds the reference to the two nodes in the different networks which are identical (or connected by ferries) in reality. While the connectivity in the network is usually assured by a node common to two adjacent links, the 'tele-porter' describes thereby connectivity between two nodes. Creating the tele-porter data objects has usually to be done manually. In case of the EuroRoadS data sets, Swe-den and Norway have supplied bordernode informa-tion provides a direct reference to the bordering data set and the node ID. Thereby, the data users's inter-nal 'tele-porter' object could be easily created.



Net-matching for reference transfer

Finally, the detail of geometrical network description was reduced ('coordinate thinning') since the data volume for larger coverages would otherwise be too big to be handled by the memory and processing capacity of the mobile device. Since the data supply provides usually equidistant coordinate pairs for each link, many coordinate pairs on links with little curva-ture can be omitted, without loosing much quality in the geometrical description of a link.

Lessons

The EuroRoadS framework – as far as tested in the demonstrator – has been experienced by the data users and data suppliers as well conceived and well documented. It offers a number of concepts e.g. for the network description or the linking of borders which appear well formed. In general and typical circumstances, the barrier to provide data in conformance with the EuroRoadS framework should be relatively low to a data supplier. This is important for the introduction of EuroRoadS.

With the EuroRoadS framework being relatively flexible and open, the harmonisation of data sets from different suppliers has therefore to occur on the data user's side. Ultimately, this burden needs to be shared by data suppliers and data users, e.g. by certain conventions on the supplier side how the EuroRoadS framework shall be used in a specific context (for a specific content type or usage).

The data volumes produced by the different test sites were quite important due to the XML/GML-exchange format. For larger coverages, this requires XML writing and parsing techniques on the supplier and the data user's side specially adapted to large data volumes.

The experiences show that useful content apart from the network (e.g. speed limits) can be transferred and integrated into a final application without full harmonisation of the road network description ('selective' use scenario). Such a scenario appears useful to many other content layers but comes at a cost. It requires techniques for reference transfer, sufficient know-how and resources for the necessary data processing on the data user's side. This step also is important with regards to data quality, the reference transfer always creates some (limited) loss in data quality.

Contact

For further information about Data Processing,
please visit EuroRoadS on **www.euroroads.org** or contact
Work Package leader for WP7 Demonstration:
Michael Landwehr, PTV AG
E-mail: michael.landwehr@ptv.de