ROad **S**afety **ATT**ributes exchange infrastructure in **E**urope

Deliverable D4.1

Description of applicable and viable data integration methods

Version Number: 1.3
Produced by: Tele Atlas
Due Date: 31/05/2009
Release Date: 01/02/2011

ROSATTE is co-financed by the European Commission – DG INFSO
Contract Number: 213467

Programme Name	7th Framework Programme – Specific Programme Cooperation – Theme 3 “Information and Communication Technologies”
Grant Agreement Number:	213467
Project acronym:	ROSATTE
Project name:	ROad Safety ATTributes exchange infrastructure in Europe
Project start and end:	January 2008 – December 2010 (36 months)
EC Project Officer:	Mr. Emilio Davila-Gonzalez E-mail: emilio.davila-gonzalez@ec.europa.eu
Dissemination level:	PU

Title of document:	Description of applicable and viable data integration methods
Work package:	WP 4
Author(s):	Stephen T’Siobbel - Tele Atlas (editor) Landwehr – PTV Ramdane Mahiou – ASFA Kees Wevers, Ahmed Nasr - NAVTEQ
Project co-ordinator:	Maxime Flament (ERTICO – ITS Europe) Tel: +32 2 400 07 35, fax: +32 2 400 07 01 E-mail: m.flament@mail.ertico.com

Abstract:
The document “Description of applicable and viable data integration methods” presents the current situation on how safety relevant data provide by road authorities and can be integrated into digital map databases.

Keyword list: digital maps, data exchange, conflation, Advanced Driver Assistance Systems

Document Control Sheet

Main author(s) or editor(s): Stephen T'Siobbel – Tele Atlas (Author and Editor)
Michael Landwehr – PTV
Ramdane Mahiou – ASFA
Kees Wevers, Ahmed Nasr – NAVTEQ

Work area: WP 4

Document title: D4.1 – Description of applicable and viable data integration methods

Version history:

Version number	Date	Main author	Summary of changes
0.1	23/09/2008	Stephen T'Siobbel Tele Atlas	First Draft – document outline
0.2	25/11/2009	Stephen T'Siobbel Tele Atlas	Revised document outline
0.3	23/02/2010	Micheal Landwehr PTV	Use case Bavaria + revision
0.4	03/03/2010	Ramdane Mahiou	Use case ASFA
0.5	04/03/2010	Stephen T'Siobbel Tele Atlas	Revisions, References, Annexes
0.6	08/03/2010	Stephen T'Siobbel	Revisions
0.7	08/03/2010	Kees Weevers NAVTEQ	Contributions by NAVTEQ
0.8	08/03/2010	Stephen T'Siobbel	Revisions
0.9	23/03/2010	A. Nasr & T'Siobbel	Revisions
1.0	30/04/2010	Stephen T'Siobbel	Final draft for EC review
1.1	28/05/2010	Stephen T'Siobbel	Consider feedback from EC reviewers
1.2	20/10/2010	Stephen T'Siobbel	Feedback EC reviewers
1.3	01/01/2011	Stephen T'Siobbel	Feedback EC reviewers

Approval:

	Name	Date
Prepared	Stephen T'Siobbel, Tele Atlas	01/04/2010
Reviewed	Dan Adams, Tele Atlas	19/04/2010
Authorised	Jean-Charles Pandazis, ERTICO	31/01/2011

Circulation:

Recipient	Date of submission
ROSATTE Consortium	01/02/2011
EC	01/02/2011

Excutive Summary

This deliverable presents results of ROSATTE WP 4 task D4.1 - Description of applicable and viable data integration methods.

The overall objective of WP 4 is to assess the fully automatic integration of aggregated safety attributes provided by public sector into pan-European seamless digital map database in a controllable and timely manner using the ROSATTE exchange infrastructure and exchange format(s). This is done by building prototypes of ROSATTE data integration service meeting the requirements formulated in WP 1 and according to the specifications formulated in WP3. The testing and validation of the complete ROSATTE chain, thus including Data Integration, is described in D5.4 – Aggregated Test Report. A revision of this deliverable is foreseen after the validation period.

D4.1 starts with a general introduction to the ROSATTE project and states the purpose of this report.

Chapter 2 gives a general introduction to data integration, the historical background of (semi-) automatic integration/conflation, and the state-of-the-art approaches followed by leading map providers.

The relevant for requirements for the ROSATTE data integration are represented and discussed in Chapter 3.

Chapter 4 addresses these requirements and discusses ROSATTE relevant data integration aspects on data collection, data versioning, data processing, location reference decoding, and data feedback.

A number of ROSATTE relevant use cases are presented in chapter 5. These explain data integration technologies from perspective of different stakeholders. The details of the software components developed for the data integration effort of ROSATTE will be represented in another report: deliverable D4.2.

Table of Contents

Index of Tables

Index of Figures

1. Project description

1.1. ROSATTE Contractual References

ROSATTE is a STREP submitted for the call FP7-ICT-2007-1. It stands *for ROad SAfety ATTributes exchange infrastructure in Europe.*

The Grant Agreement number is 213467 and project duration is 36 months, effective from 01 January 2008 until December 2010. It is a contract with the European Commission, DG INFSO.

The EC Project Officer is:

Emilio Davila-Gonzalez
EUROPEAN COMMISSION
DG INFSO.G.4 - ICT for Transport
Office BU31 4/50
B-1049 Brussels
Tel: +32 2 296 21 88
E-mail: Emilio.Davila-Gonzalez@ec.europa.eu

1.2. Project Objectives and scope

The ROSATTE project intends to develop the enabling infrastructure and supporting tools that will ensure European access to road safety attributes including incremental updates. This infrastructure will facilitate administrative internal functions as well as supply of data to third parties e.g. for safety relevant services.

The overall objectives of the project are to:

1. Facilitate access to, exchange and maintain European-wide core road safety spatial data from national/regional/local sources by standard procedures
2. Enable multi-level aggregation and update of European-wide safety map data
3. Assess the technical and organizational feasibility of this infrastructure

1.3. Key issues / Project scope

Accurate and up-to-date safety related road network attributes are particularly important for efficient road operation and administration, and for safe driving along the European road network.

For data users, the reality today is, however, a rather complex landscape of multiple data providers, multiple formats, varying availability and quality of data and long delays between data updates. Road authorities and infrastructure operators are usually at the beginning of the information chain, being responsible for the planning, change approval, physical implementation, equipment and maintenance of roads.

The major problem is how to ensure timely and easy access to road information owned and maintained by thousands of road authorities. In addition, mechanisms are needed to enhance the quality of the available data in terms of accuracy, correctness and up-to-dateness; and, to enable multi-level (local/national/European) aggregation of the data. With respect to a future continuous delivery and integration of updates of road attribute data, road authorities that provide such updates will be responsible for the timeliness of delivery (within an agreed time period after the change of the attribute on the road), and for the correctness and positional accuracy of the data. Data integrators will be responsible for correct interpretation of the received data, and the correct inclusion in their digital map databases. For specific safety critical attributes, an independent certification body may be created that will be responsible for surveillance of the methods and procedures used.

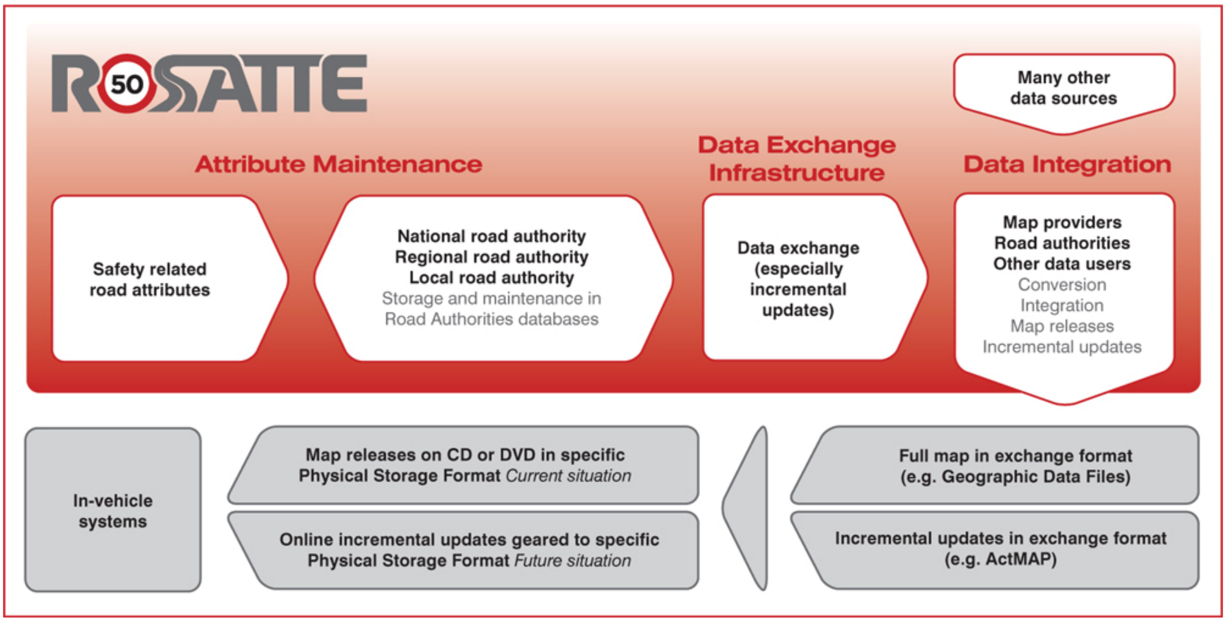


Figure 1 - The scope of the ROSATTE project

1.4. ROSATTE Project deliverables available to date

Since the start of the ROSATTE project in 2008, a number of reports have been finalized. The reports labeled for public dissemination are - or will be - available via the ROSATTE project website (www.rosatte.eu). The table below gives a comprehensive overview of available reports to date.

D1.1 State of the Art
Describes the current road authorities and infrastructure operator's situation with respect to how safety relevant data is stored, exchanged and updated.
D1.2 Requirements and Overall Architecture

Defines the project scope, user, user requirements and derived system requirements. It also gives suggestions on information model and a high-level system structure.
D2.1 Specification of tools
Conceptual specification of a road data maintenance & delivery system
D2.2 Implementations of tools
Implementations of tools for demonstration of data maintenance and access in different test beds
D3.1 Specification of Data Exchange Methods
Presents the exchange specification consisting of a data content specification, a physical exchange format specification, (to be completed) and a service specification.
D5.1 Test and Validation Plan
Describes the processes for validating the objectives of the project on each test site, tackling quality aspects, user requirements and evaluation methods.

Table 1 - Overview of project deliverables available on http://www.ertico.com/en/activities/safemobility/rosatte/publications/publications.htm to date.

1.5. Purpose of this document

The overall objective of the Data Integration work package (WP 4) of the ROSATTE project is to assess the possibilities of automatic integration of safety attributes provided by the public sector into a pan-European seamless digital map database in a controllable and timely manner using the ROSATTE exchange infrastructure and exchange formats. This work package analyses, develops, tests, and validates one or more implementations for the integration of the safety attributes.

This document is addressed to all stakeholders in the ROSATTE safety data provision chain: public (road) authorities, service providers, map providers, private providers of safety attributes, vehicle (navigation) system companies, vehicle companies, etc.

From the ROSATTE description of work, we learn that this specific deliverable (D4.1) addresses the first task which was identified in this work package:

Data integration/aggregation technologies assessment (Task 4.1), discussing the:

- Identification and review of viable candidate technologies for the information decoding, data matching, and data integration based upon the practical data exchange technologies using different location referencing mechanisms proposed in WP3.
- Selection of technology(ies) for reference and test site implementations.

However, the concepts for ROSATTE which emerged during the first part of the project (e.g. ROSATTE Data Store specification) and the results of the work performed in work package 3 (describing the detailed specification of web services to transfer information, the selection of location referencing technique AGORA-C [2, 3, 4], shed new light on the

tasks that were identified in the original description of work of the ROSATTE project. This will be described in chapter 4 of this report.

For a complete view on work package 4, the other tasks include:

- *Quality assured data integration (Task 4.2)* which analyses the impact on quality of the used data decoding, data matching, and data integration methods and tools, and defining of automatic and semi-automatic quality check routines which can be integrated in existing procedures.
- *Design and implementation of data integration/aggregation technologies (Task 4.3)* which designs and develops software components including quality check routines applicable to integrate information from the ROSATTE test sites.
- *Implementation aspects (Task 4.4)* which focuses on the organizational, architectural and quality aspects of data aggregation and integration.

The description of the developed software components dealing with the data integration for the ROSATTE project will be available in deliverable D4.2.

2. Integration of Spatial Data

2.1. General Introduction

Road databases need to be kept up-to-date to reflect the changes in the real world they represent. The quality of such databases to a large extend depends upon how successful the map providers are in maintaining their databases. With the ever increasing spatial coverage of these maps (e.g. pan European, North America/Canada) and with ever more stringent quality requirements from the map users, maintaining & improving up-to-dateness represents an enormous challenge. Cost efficient and timely processes for digital map maintenance need to be in place to meet the user expectations. This is particularly true when digital maps are to play a role as a "predictive sensor" in a vehicle, often complementing information retrieved from other sensors (camera, radar, laser) in order to support or enable in-vehicle safety applications (e.g. speed alert, lane keeping, curve warning system) [19].

In the past, map providers relied upon field survey to inspect the correctness of the digital maps and make updates whenever necessary. By introducing mobile mapping, vehicles equipped with multiple cameras and positioning sensors logging information along the travelled roads, on a large scale, the actual inspection and updating of the map can be done in a (semi-) automatic way, and at remote locations. The key aspect here was to split map data collection from data processing, and by doing so, make the process more scalable. Teams of many digital cartographers can therefore process information collected by one survey vehicle, wherever in the world [27].

Other processes which were developed in the past to enrich, extend, and maintain spatial databases address the consumption of other (spatial) databases. The concept is to inherit from other geospatial & non geospatial databases those properties, features and attributes which are relevant to improve the quality of a core spatial database. The latter forms the base map where other map products are derived from, e.g. maps for navigation, maps for Geographic Information Systems (GIS), maps for Location Based Services (LBS), etc. The term "conflation" is widely used to refer to this integration of data from different sources. Without the ability to conflate/integrate data from many different sources, data users are faced with duplication of effort and related unnecessary cost. In this document the terms integration and conflation are used interchangeably.

In GIS, the conflation process combines the information from at least two spatial data sets to make a new composite data set that is superior to the input source data sets in either spatial or attribute aspect. The objectives of conflation include: increasing spatial accuracy and consistency, updating or adding new spatial features into data sets; updating or adding more attributes that associate with the spatial features of the data sets, etc. The spatial data sets in conflation are defined as digital map databases or digital spatial files that cover the same area, describe the same information in different forms and vary in density and accuracy [5]. Conflation also can add new attribute information from one data set that does not exist in the other. Adjacent maps may be integrated to a single map by a process called edge mapping. The following paragraphs gives insight in the essentials of the challenges related to geographic data set integration.

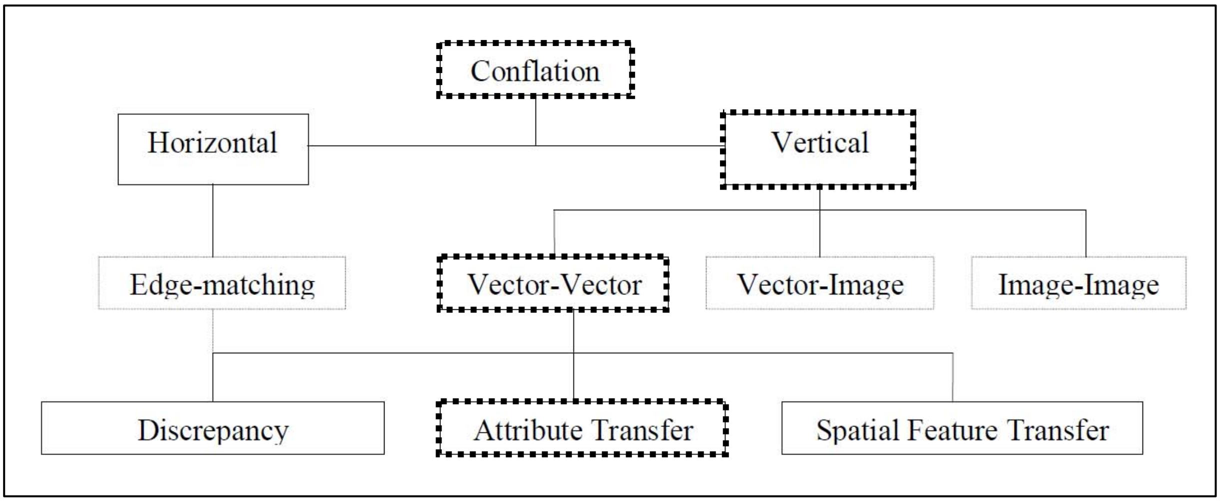


Figure 2 - General conflation classification (Yuan and Tao, 1999). Applicable parts for ROSATTE are dashed.

Yuan and Tao (1999) defined two types of map conflation, horizontal and vertical conflation, see Figure 2. A third type was added by Vivid Solutions [29] in their JCS conflation suite: internal conflation. The differences between these classifications of conflation are:

- Horizontal Conflation is referred to edge-matching of neighbour maps with the objective to eliminate spatial feature position and attribute discrepancies which exist in the common area of the two vector or raster maps. Examples are aligning the boundaries of adjacent coverages, or edge-matching neighbouring networks. These kinds of operations are also known as boundary alignment.

- Vertical Conflation involves matching and/or eliminating discrepancies between datasets that occupy the same spatial area. Examples include road network matching between two representations of roads in the same region. Two important kinds of vertical conflation are Version Matching and Feature Alignment. In the case of Version Matching the input datasets consist of different versions of the same features. The conflation process is intended to identify matching features. Attributes may be transferred between matched features, and unmatched features may be transferred in their entirety. An example of this is matching different versions of road networks for the same geographical area. In case of Feature Alignment the input data consists of features from two or more different feature classes that bear some defined relationship to each other. A common relationship is that of geometric alignment of administrative boundary features. An example of this is aligning the boundaries of different kinds of feature classes such as municipal districts and lot parcels.

- Internal Conflation, this kind of conflation is only in use on single data sets, and is used in the process of removing overlaps, cleaning of coverages and quality assurance.

The existing conflation technologies can be categorized into three groups based on the spatial datasets they deal with:

- Vector to vector data conflation. E.g. conflation of two road networks of different accuracy levels.
- Vector/raster and raster/vector data conflation. E.g. alignment of a road network to an aerial photograph of the corresponding area, possibly followed by annotating the raster image; or integrating altitude data from a raster terrain model into a road network draped over this terrain data.
- Raster to raster conflation. E.g. aerial image and a raster map to improve visual appeal.

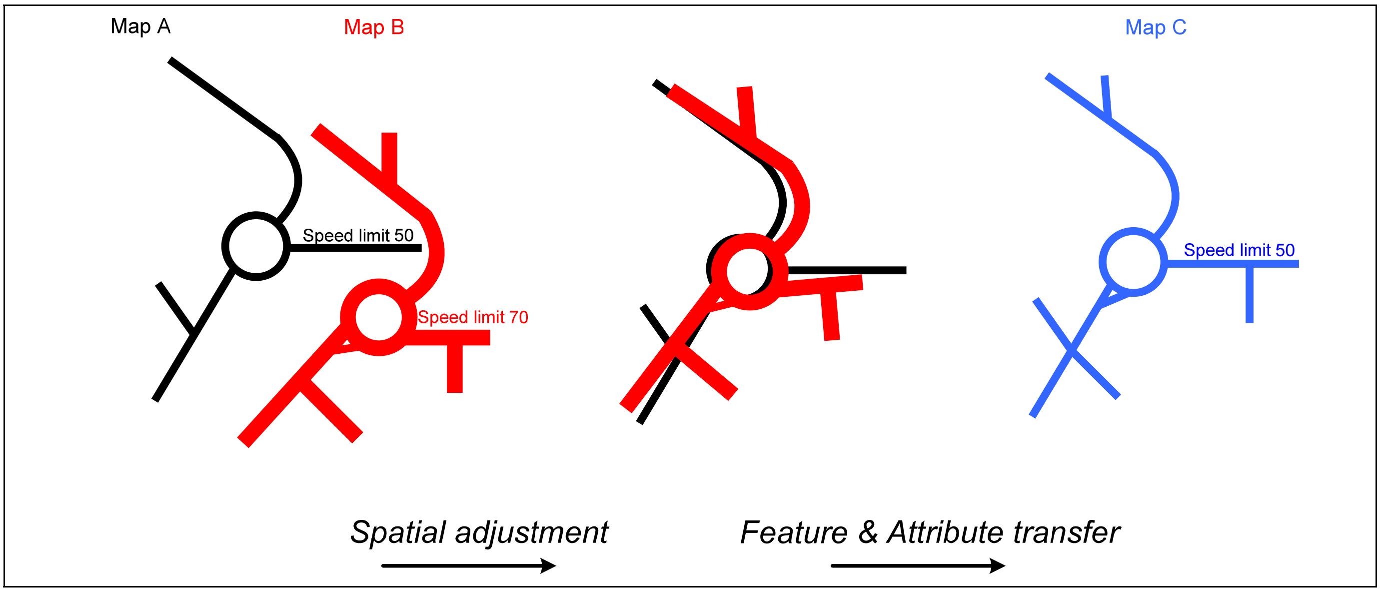


Figure 3 - Example of vector-vector map conflation where features (new roads) and attributes (speed limit) are conflated.

Conflation can be used to solve different practical spatial problems. Examples of conflation applications include:

- Spatial discrepancy elimination. A global adjustment of spatial feature coordinates in one or both data sets so as to eliminate the feature position discrepancies such as sliver polygons, shifts of features, etc. This application exists in map edge-matching, map compilation, etc.
- Spatial feature transfer. By recognizing the common features of different maps and identifying those not matched, new features can be added into the old map, or old coordinates can be updated. This application mainly exists in GIS spatial data updating.

- Attribute transfer. This step may be considered a relatively trivial task if a one-to-one correspondence exists between the source and destination datasets, but such ideal cases are rare. In the real world, many scenarios arise. For example, one feature in the source may have many corresponding features in the destination dataset, or many features in the original can have a single corresponding feature in the destination. To deal with a many-to-one match in the traditional conflation approach, the feature in the destination needs to be split into smaller segments in order to create several one-to-one matches, thus allowing the desired attributes of the source to be transferred. In the one-to-many match, some attributes, such as address range in the source dataset, may need to be split and then transferred to target features. Both spatial and attribute splitting have to be done in many-to-many matching cases. The manual splitting operation is time consuming. In addition, many road datasets may have a route system defined in terms of its existing segments. The splitting of segments would likely corrupt any defined route system. Manually regenerating route systems after conflation is undesirable. (from Song et al., 2006)

2.2. Challenges for geographic data set integration

There are two broad categories of factors responsible for the differences between geographic data sets (from [28]):

- Differences in contents. Geographic data sets are collected for specific purposes, sometimes totally different from one set to the next one (different themes). It is a representation of a set of real-world phenomena. Different sets of real-world phenomena will imply different contents among geographic data sets.

- Differences in abstraction and level of detail. In capturing real-world phenomena there is the process of transforming real-world phenomena into a data set representation. Different rules for surveying, for the same terrain situation, may lead to different object classes, with different attributes, and different geometric descriptions, by points, lines, or polygons. Above all, these differences make it important to develop an understanding of the semantics of the data sets. Semantics should be understood as the link between a real world situation and a data set representation.

Spatial database integration, according to Devogele et al. (1998), is the process of integrating more than one heterogeneous and autonomous spatial data set into a single unified description of reality.

2.3. History of (semi) automated spatial data integration

In this chapter, the wide scope of "spatial data integration" is narrowed down to reflect situations of vector-vector conflation in the GIS world, and even to road database conflation.

The history of automated map conflation goes back to the early 1980s, when the first development and application of an automated conflation process occurred during a joint

U. S. Geological Survey--U.S. Census Bureau project designed to integrate the agencies' respective digital map files of U.S. metropolitan areas [24]. The major concern of conflation then was to eliminate the spatial inconsistency and improve the spatial accuracy of digital maps. When two vector datasets are spatially aligned, attributes can be transferred from one map to another, or the two maps can be fused together to one. By performing this data fusion, the best properties of each map can be inherited to the remaining map database. It is recognised that the implementation of a computerized system for this task provided an essential foundation for much of the theory and many of the techniques used today [7].

To automatically collect and update GIS databases, Walter and Fritsch (1999) proposed a relational matching approach to find matched spatial objects based on the similarity of spatial objects at the geometry level (e.g., node to node matching based on distance) and based on the relations between the elements in a dataset. They investigated the similarity of spatial objects based on statistical information derived from a random sample of the vector datasets to be integrated. The developed method was tested on road networks and matched geometric elements in two different databases defined at similar scales.

In addition to performing feature matching at the geometry level, Cobb et al. (1998) proposed an approach to perform feature matching at the object level. For example, when comparing two road segments, their approach not only matches the road endpoints, but also matches the non-spatial properties such as street names and widths. Ware et al. (1998) presented a technique for matching and aligning vector features in pairs of multi-date coverages. The main feature dealt with in their work is road segments. Their approach also supports the additions/deletions of road nodes in order to deform the matched road segments. More generally, many matching algorithms for two network databases at similar scale [30], [34], [12] or different scale [10], [35] were proposed.

Indicative of the importance of conflation/data integration, is the availability of dedicated tools in popular GIS toolkits. In fact, there are several commercial products available for the ArcGIS platform such as MapMerger [18] and Conflex [9] that support automatic vector to vector matching with limited human intervention to consolidate multiple vector datasets.

Over the last years, also Open Source tools to conflate spatial data have emerged. Examples include JUMP [17], JTS topology suite and JCS Conflation Suite [29]. The latter is implemented as a number of Java plug-ins for the JUMP Unified Mapping Platform.

Leading map providers have since the start of their activities been working intensively on conflation techniques to extend the coverage of their digital maps, enhance their content and improve their up-to-dateness. Over the years, this has lead to the existence of a suite of proprietary conflation tools and techniques which can be tailored to retrieve information from available spatial data sources [34]. Often, the developed technologies require some level of human intervention as high quality of the result is of utmost importance.

An excellent example - and probably the largest digital map conflation project ever undertaken - is the conflation of the Geographic Data Technologies (GDT) North American map database with the Tele Atlas North American map database after map provider GDT was acquired by latter map provider in 2004. This resulted in a digital map which inherited the best features and attributes of the two map databases. Generalized, the resulting geographical database used the geometry advantage in the database and attribution advantage with the GDT dataset.

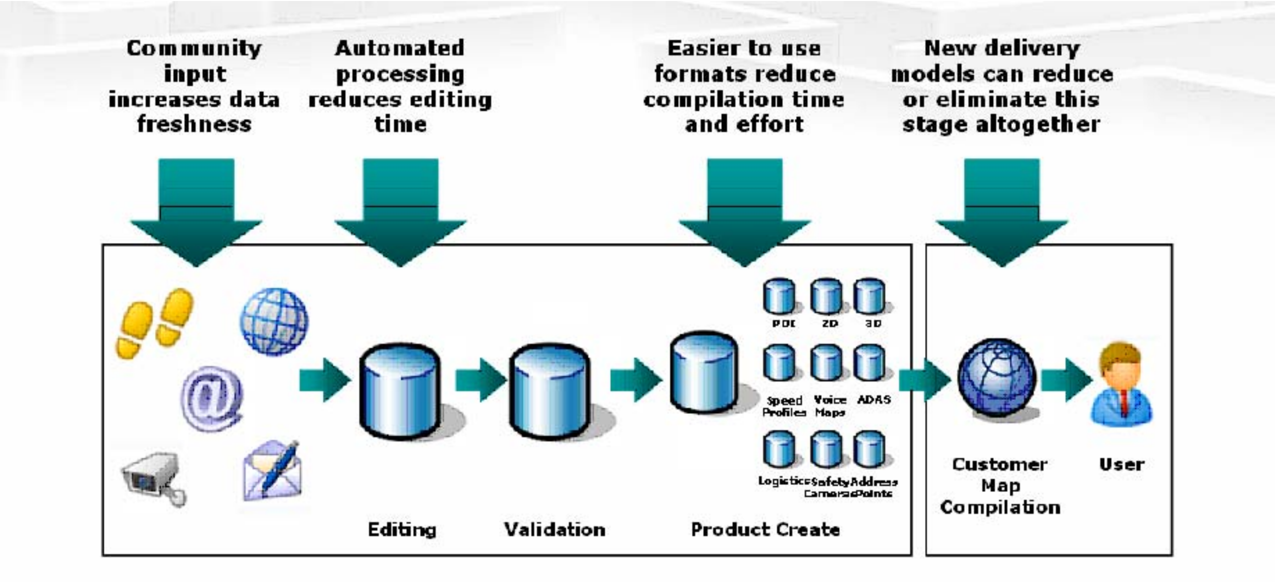


Figure 4 - High level overview of state-of-the-art digital map production and maintenance steps. Source data is available via field survey, community input (map users reporting deviations, and passive feedback in the form of probe data, other (spatial) data providers including governments.

2.4. Recent approaches in digital map creation and maintenance: web reporting tools and feedback from navigation system users

The last 5 years we have seen web-based reporting systems emerging which collect information about differences between map providers data and the real world, such as NAVTEQ Map Reporter. This approach is just one of many inputs into the map provider's operational systems, in addition to other existing sources. The output of these reporting web-tools helps directing map providers where to deploy resources to research and analyze changes in the road system. Map providers generally do not automatically update the map data just because a report was entered into the reporting tools. When map feedback is entered into the tool, the system analyzes many aspects of the issue being reported and how it relates to the rest of the data. For example, if there is a large number of addressing issues in one area, it might indicate a change in the street geometry rather than a series of relatively minor addressing issues. Recoding each minor issue individually would be costly and wouldn't address the larger question of street geometry. Thus the reporting tool enables map providers to make more efficient use of their data editing resources. Today, the updates which can be provided by map users do not cover all the ROSATTE safety attributes. However, this is expected to change in the near or mid term.

NAVTEQ Map Reporter makes submitting a map report (or database update request) faster, easier and more efficient for customers and consumers. Reports are submitted via a map-based interface on the web site. Each submitted report is followed up individually and "map reporters" are able to monitor the resolution of their report and if a change is being made to the database. Reports identifying a legitimate map database discrepancy are investigated and corrected in the map provider database. Customers submitting reports can check on the status of their submissions by running customized reports using criteria such as submission date, user, status, or country location. Map Reporter combines the benefits of community generated content and feedback with the strengths and capabilities of NAVTEQ, to deliver consistent, reliable maps and broad coverage. Figure 5 provides a screenshot of the Map Reporter web interface.

One of the major advantages of this web reporting approach is that the respective updates are directly referenced according to the map provider map database. Spatial conflation of information is therefore relatively straightforward. As the map updates are provided by users of a specific map, there is no exchange of these updates between map providers. In some cases this might not be necessary as the map of another map provider might already be up to date.

A second approach of reporting updates for maps, such as MapShare, allows drivers to make changes to the maps directly on their navigation devices and share them with others. These updates can be shared with other map users. Users who connect their devices to their computers can download and upload map corrections on a regular basis via a software tool. The users are able to select various 'levels of trust', e.g. only receiving changes that have been verified by the map provider or changes that have been submitted by trusted sources (such as road authorities).

The advantage of web reporting and driver feedback approaches is that map updates are spatially referenced to the same map that is in use by the person reporting the update. This eliminates possible spatial mismatches when different maps are being used during the map update process at the map provider's side.

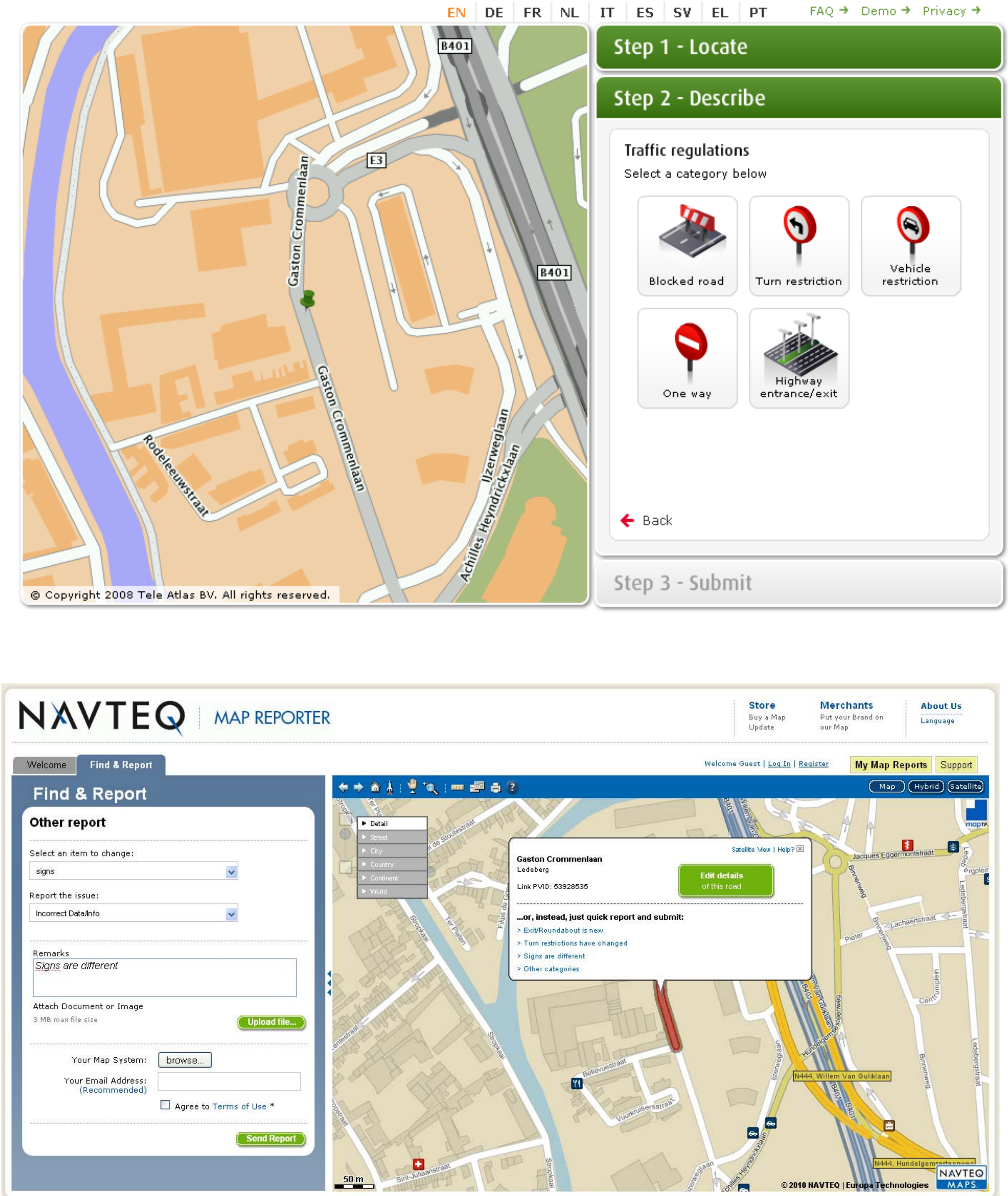


Figure 5 - Screen shots of web browser tools by Tele Atlas (top) and NAVTEQ (bottom) allowing people to report improvements to the available digital map databases via the internet.

2.5. Spatial data Integration in steps

The workflow in a semi-automatic conflation process generally can be broken down into the following common subtasks (overview from [6]). Depending on the nature and quality of the data in a specific conflation problem some of these subtasks may be trivial or not required. The conflation tasks below typically reflect:

- Data Pre-processing. This step normalizes the input datasets to ensure that they are compatible. For instance, they must have the same coordinate system. This may also involve format translation and any other basic preparation of the datasets.

- Data Quality Assurance. Some conflation tasks require that datasets have a given level of internal consistency. For instance, coverage alignment algorithms require that the input datasets are in fact a clean coverage. During this step the internal consistency of the datasets is verified and if necessary improved.

- Dataset Alignment. In some cases, datasets are sufficiently misaligned that an initial alignment process is required to allow more precise conflation to be carried out. This alignment is typically coarse grained in nature, not descending to the level of aligning individual features. This process may be either manual or automatic.

- Feature Matching. During this step common features between the datasets are matched. This may be done either in an automated fashion using one or more conflation algorithms, or via manually determined matches. After this phase has been performed the discrepancies between the datasets will have been identified. It is often useful to provide statistical summaries of data quality, or to visualize the discrepancies.

- Geometry Alignment and/or Information Transfer. Once map features have been matched, the match information can be used to improve the quality of one or both of the input datasets (correctness, up-to-dateness, accuracy, …)

- Geometry Alignment removes discrepancies between geometries.

- Information Transfer involves updating one dataset with information from the other. This information can be either attributes or geometry to be added to an existing feature, or entire features to be added to the dataset.

- Quality Checking. This is often performed in a semi-automatic way. With more and more up-to-date and possibly redundant information becoming available (data from vehicle probes, active feedback from navigation device users, etc.) performing quality checks can be facilitated.

2.6. Location Referencing

Modern telematics systems bear a potential to avoid traffic jams; especially car navigation systems can support public traffic management to redirect traffic flow in order to relieve overloaded roads. The need for systems to collect, gather, distribute and process traffic information for geographic areas arises in many parts of the world. Various systems worldwide are already in use and in many countries the installation of such a system is in progress. Apart from traffic jams, there is an increasing demand for communicating other kinds of georeferenced information, such as:

- Traffic Information in general,
- Emergency/Accident Location,
- Yellow Pages,
- Location-based Services, and
- Points of Interest.

Hence, we have to deal with the task of georeferencing or how to re-identify certain designated places. This is particularly more difficult when the georeferenced location needs to be re-identified in the context of a different version of the map or a version from a different vendor.

Compared to georeferencing in general, location referencing is the restriction to identifying elements of a road network. It does not identify other objects like rivers, train tracks, border of lakes or forest etc.

A ROSATTE service is not designed to be a real time service e.g. to support Variable Message Signs. However, with an expected up-to-dateness of 24 hours, and the fact that an update set of safety attribute(s) may contain just one single attribute update with a particular location reference, or a set of location references, a suitable location referencing technology is of major importance to ROSATTE.

In the past the location referencing task was solved by assigning static codes to a road sections. These codes are maintained in tables and have to be implemented on the service provider side, in the digital maps and the end user terminals. Due to the bandwidth capacity and the memory restrictions of the end user terminals the number of codes have be small. The maintenance of these codes are quite expensive and there is always a time gap between releasing new tables with codes and their market penetration. Besides that you have to take care about backwards compatibility. This form of static coding is used by TMC, VICS, etc. method. To overcome these disadvantages AGORA-C was developed. This AGORA-C emerged from a number of initiatives on European scale, such as ERTICO Location Referencing Committee [37] and EVIDENCE [38].

The AGORA-C method unifies a number of earlier, proprietary methods which – if applied on their own – perform less reliably and robustly. Each of these earlier methods emphasized on a certain characteristic of geographical maps, namely road shape, road connectivity, cross road & side road information, or road object classification. By combining these methods, AGORA-C showed to be superior as a more comprehensive and selectively redundant approach.

During the final stage of standardizing AGORA-C in ISO/TC204, elements from another dynamic location referencing approach have been incorporated, namely the results from

Japanese research activities led by Panasonic. The Japanese method was strongly based on road shape and maximized compression to meet unique Japanese conditions (lack of road names that can be used as descriptors, and bandwidth constraints of the current communications/broadcast infrastructure).

Parallel to the initial phase of AGORA-C developments, another proprietary method for dynamic location referencing has been designed and evaluated by DaimlerChrysler with support from Tele Atlas and NAVTEQ. This method, referred to as ROSA, is constructing a composite reference (road intersection as a primary (absolute) location and a traffic event or Point Of Interest as a secondary (relative) location) and relies on a classification scheme for geographic objects. Opposed to AGORA-C which only prescribed the encoding rules, ROSA also prescribes how to decode by means of a correlation function. Due to the fact that this method is using fewer geographic characteristics than AGORA-C (hence leading to lower matching rates) and that it may be affected by interpretation ambiguities in regards to object classification, ROSA developments have been discontinued in 2003.

In the area of data services via digital broadcast media, a data streaming protocol called TPEG (Transport Protocol Expert Group) [39] has been developed for providing rich information presentation to digital radio receivers, including any mobile computing devices with broadcast reception capabilities, as well as internet browsers. Mobile devices may or may not be GPS enabled, and may or may not use an on-board digital map. The range of so-called TPEG applications includes road traffic messages, congestion and travel times, parking information, and public transport information. In order to support location-sensitive filtering and presentation (such as viewing events in the user's proximity), a native location referencing method has been developed. This method, TPEG-Loc, has some commonalities with some of the roots of AGORA-C (i.e. cross road and side road information). TPEG-Loc and AGORA-C are not directly competing technologies, because they are designed for different use cases. AGORA-C is designed to support "machine-to-machine" location descriptions optimized for size efficiency and autonomous decoding. TPEG-Loc provides for "journalistic" location descriptions in terms of a structure and grammar for location classifiers (including coordinates) and full text descriptors. In comparison, TPEG-Loc is a more general approach than AGORA-C and less specialized for a particular domain (automotive) and a particular class of client devices (dynamic route guidance systems and location based services). TPEG-Loc does not provide for standardized encoding rules, nor does it provide for means of compression for optimized encoding compactness. In essence, use of TPEG-Loc location references will on average consume more storage space or bandwidth and will systematically result in a lower matching rate.

The most recent development in Location Referencing is OpenLR™ [20]. OpenLR™is an open source software project launched by TomTom International B.V. in September 2009. The project provides a royalty-free dynamic location referencing method which enables reliable data exchange and cross-referencing using digital maps of different vendors and versions. Thus, OpenLR™ will help to enhance existing applications and will generate opportunities for new services.

Since OpenLR™ is open source everyone is invited to contribute to and to enhance the exististing solution. The www.openlr.org website offers a reference implementation of the OpenLR method and also provides a complete documentation in terms of software documentation and a technical whitepaper. A detailed presentation on OpenLR™ also addresses both, technical and business issues.

Relevant for the ROSATTE project, in May 2010, OpenLR™ was extended to (1) support more accurate description of locations by improvements to how offsets are described and (2) support point locations.

3. ROSATTE Requirements

3.1. Requirements from D1.2

A comprehensive overview of the ROSATTE system and user requirements was presented in deliverable D1.2 of the ROSATTE project [22]. The system requirement tables, which reflect functional, non functional, and conformance (with European law, directives) requirements, are available in Annex 1 in this document, see Table 2.

A number of functional requirements (FR-1, FR-2, FR-3) address issues related to system for the sourcing of ROSATTE's safety attributes. These are important, but for the "integration" of safety attributes into map databases of the map providers, it is assumed that all required information was received by map providers and is available for processing, either representing an incremental update or full dataset set of updates (FR-5). Non-functional requirements reflect quality characteristics and are relevant for WP 4.

The actual integration of the provided safety attribute(s) into the map providers spatial databases starts after successfully resolving the geometry for which the updated is intended by decoding the given location reference on the receiving map (FR-6). Suitable tools are required to perform the data integration (FR-12). The result of the data integration step is of interest to those who provided updates to be integrated ("Data Store Operator" and "Enacting Authorities", see D1.2 [22]), therefore it was indicated that a feedback mechanism must be in place (FR-11). The data integration methods need to be able to handle flexible type definitions as "Road safety attributes" are not finally decided (FR-13).

The user requirements from map provider's perspective are of course evidently relevant for the selection of relevant data integration technologies, see Annex 1 - Table 3.

The following chapter presents how these requirements are interpreted and describes the ROSATTE data integration. For each aspect related to data integration, the corresponding ROSATTE requirements are listed.

4. Data integration within the ROSATTE scope

Whereas previous chapters and paragraphs presented the history as well as a general description of (semi-) automatic map conflation and all the relevant separate subtasks, this chapter will address the ROSATTE specific aspects of "data integration". This activity describes the role of the "Information Provider": the instance responsible for the provision of maps and datasets for use in ADAS applications, and the receiver of road safety attributes from data services. These are the map providers, but also the enacting authorities.

The high level concepts which have emerged during the first part of the ROSATTE project (e.g. one ROSATTE Data Store) and especially the results of the work performed in work package 3 (e.g. selection of web service for information exchange in line with the INSPIRE directive, selection of AGORA-C [2,3,4] as location referencing technology, etc.), sheds a new light on the tasks that were identified in the original description of work of the ROSATTE project:

Task 4.1 - Data integration/aggregation technologies assessment

- *Identify and review viable candidate technologies for the information decoding, data matching, and data integration based upon the practical data exchange technologies using different location referencing mechanisms proposed in WP3.*
- *Selection of technology(ies) for reference and test site implementations.*

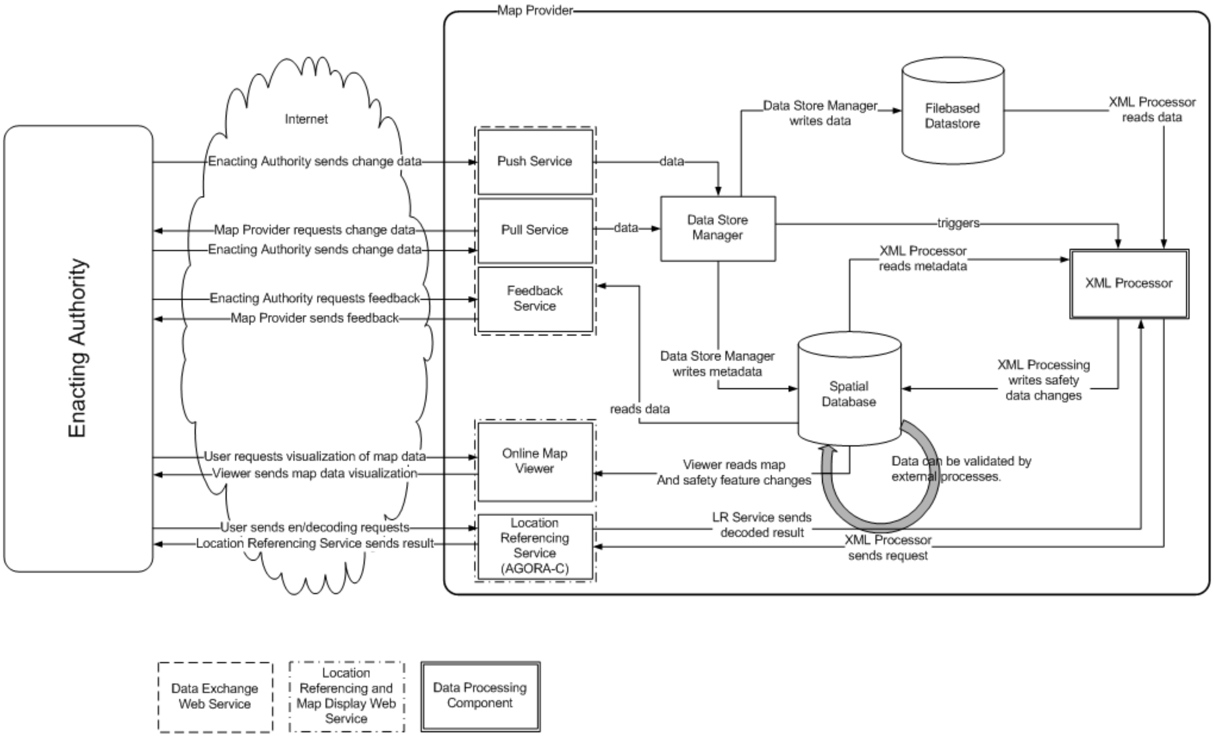


Figure 6 - Overview of ROSATTE services and processes for data integration.

In fact, we recognize that the number of “candidate technologies” we had envisaged to investigate has been narrowed down due to specific choices that were made for the definition of the ROSATTE exchange infrastructure, see WP 3. Below, an enumeration is given of the different processes & technologies and their specific relevance for the ROSATTE Data Integration.

4.1. ROSATTE data provision to information providers (map providers)

- Options: ROSATTE safety attributes may be provided via media as CDs, DVDs, Hard disks, over the internet via FTP and dedicated web services.
- ROSATTE requirements: List of requirement to comply with INSPIRE.
- ROSATTE selection: Use web services, using Representational State Transfer (REST) [21] or SOAP [25] see also D3.1.
- Review: The use of CDs, DVD, Hard disks, etc. could still be relevant in the context of a ROSATTE provision of safety updates for “initial deliveries”. These may represent bulk deliveries of vast amounts of information typically for larger datasets reflecting countries or provinces. For example, such “initial delivery” may describe a complete set of ROSATTE safety attributes represented as “traffic signs” for the province of Flanders in Belgium. This information may be rather voluminous due to its abundance, coverage and its representation (XML). Although data compression techniques may be applied successfully, it may still be less convenient to use web service for example. Authorities, bodies, companies, etc., providing ROSATTE safety attributes could decide to provide a yearly, bi-yearly, etc. “initial update”. This is comparable with incremental map updating concepts where frequent smaller updates refer to base maps which are issued e.g. on a yearly basis (e.g. ActMAP [1]). Of course, by distributing updates & update dataset via CD, DVD, and Hard Disks, it will most likely not be possible to achieve the targeted maximum 24h detail between the provision of the update and its successful integration in the map at the map providers’ side (see WP 5, deliverable D5.1). For the fast transmission of individual safety attributes updates or smaller update datasets, the use of the internet is recommendable. In fact, the specification work for the exchange infrastructure WP3, describes the use of RESTfull web services [21].

4.2. ROSATTE data versioning

A versioning scheme has to be defined to allow the clients and servers which are part of the ROSATTE infrastructure to agree about the content of a specific update data set. When these update data sets are retrieved by the map providers and wait for further processing, the subsequent data versioning is dependent upon the internal processes of the map providers.

- Options: For identifiers, use can be made of serial numbers, random numbers, names or codes allocated by choice which are forced to be unique by keeping a central registry such as the EPC Information Services, the Open Software Foundations (OSF) standard for a Universally Unique Identifier (UUID) [33].
- ROSATTE requirements: The versioning scheme has to be flexible enough to allow the enacting authorities to define update intervals which suit their needs and allows to keep up their specific workflows when managing safety attribute related data.

- ROSATTE selection: From D3.1: The versioning scheme of ROSATTE defines how the ROSATTE "DatasetId" is generated which uniquely identifies each dataset which is made publicly available by an enacting authority. The DatasetId is represented by two 128 bit values where the first consists of an OSF/DCE UUID value identifying the enacting authority and the second consists of a high 64 bit part and a low 64 bit part. The high 64 bit part represents the start date of the update interval – the low part of represents the end date. Each of the 64 bit values contains the Unix time of the specific start- or end date. The Unix time is widely used and can therefore be regarded as a standard which can be easily adopted by ROSATTE implementations. Note that the Unix time is defined to be the number of seconds since 0:00:00 January 1st, 1970 (UTC) – not including leap seconds. This time definition can also be found in the Java programming language and can easily be achieved using the C# programming language (see also D3.1).
- Review: The benefit of UUID is to enable distributed systems to uniquely identify information without significant central coordination. Anyone who created an UUID can use it to identify something with reasonable confidence that the identifier will never be unintentionally used by anyone for anything else. Information labeled with UUIDs can therefore be later combined into a single database without needing to resolve name conflicts. The most widespread use of this standard is in Microsoft's Globally Unique Identifiers (GUIDs). Other significant uses include Linux's ext2/ext3 file system, LUKS encrypted partitions, GNOME, KDE, and Mac OS X, all of which use implementations derived from the UUID library found in the e2fsprogs package .

4.3. ROSATTE Data processing & integration by information providers

- Options: Conflation approaches for spatial data as described in chapter 2.5 – including data pre-processing, data quality assurance, dataset alignment, feature matching, geometry alignment, information transfer and quality checking. For geometry alignment, please turn to chapter on Location Referencing.
- ROSATTE requirements: FR-5 (incremental updates), FR-6 (location referencing), FR-12 (tools for data integration), FR-11 (collection of information to support feedback loop), FR-13 (Flexible type definition). Non functional requirements reflect quality characteristics or aspects and are relevant too (NFR 1- 7).
- ROSATTE selection: ROSATTE data content was specified by ROSATTE partners, a geospatial databases was used to represent the ROSATTE update data.
- Review: Conceptual data content specification and physical format for ROSATTE has been developed according to ISO 19109, see D3.1, meaning that the data which is received by map providers is know upfront and not that different from the map providers conceptual data content specs of their core map databases. This greatly simplifies the components and algorithms needed for the actual integration of data integration.

  Figure 6 illustrates the data integration architecture as implemented by the map providers. The different services, processes, and databases result in an update of attributes of on line spatial database. Details of this will be presented in D4.2. The approach followed is to match the ROSATTE update to the map database by performing the location decoding, then populate a table of a spatial database to store the ROSATTE update information. For instance, all validation processes the map provider has in place today, may run on the ROSATTE updates. For example, it could be that a ROSATTE update of a speed limit was in fact already conducted, or that it is somewhere in the pipeline, or that it even is overruled by information

received via another source. One of the complexities that map providers have to address is the mapping of point feature information to linear feature information. For example, the customers of map providers expect speed limit information to be represented as an attribute of a linear feature (a road element). At least one of the ROSATTE partners however will provide (changes of) speed limit traffic signs attributes (e.g. Flanders).

4.4. ROSATTE Location Referencing

- Methods: Map id's (of map line, point, area features), geographical coordinates, linear references (Km markings), fixed location codes (TMC), map agnostic location references (AGORA-C [2,3,4], OpenLR [20])
- ROSATTE requirements: UR-5, NFR-3, NFR-6, see Annex 1
- ROSATTE selection: AGORA-C (binary code, XML)
- Review: The safety attribute information to be exchanged is related to real world locations on the road network, and needs to be referenced to this network. For this a robust location referencing method is needed, which permits encoding, transfer and decoding of the location information with a high degree of certainty. Methods based on proprietary link identifiers, linear referencing methods using proprietary routes/reference points, pre-coded locations not reflecting the complete road network are not recommendable. On the contrary, as any part of the network should be addressable, the location referencing method to be used should be flexible and dynamic. For this reason, map-based dynamic location referencing is the preferred approach. Currently one of the best available method for this is the AGORA-C location referencing method [2], which is incorporated in the ISO International Standard "Location referencing for geographic databases", Part 3 "Dynamic Location References (Dynamic Profile)" [4]. In tests, this method has shown a hit rate (percentage correctly matched plus percentage correctly identified as not present in the map database of the receiving side) of > 98% [3].

  Parts of the AGORA method are protected by IPR, and licensing conditions for the method have been published but they are not sufficiently clarified for typical ROSATTE type of implementation. Due to this and dependent of the actual future licensing conditions, wide-spread adoption of the method may be severely held back. This may imply that a future operational implementation of the ROSATTE approach may need an alternative method for location referencing. Despite this situation, for the initial ROSATTE research activities the AGORA-C method has been judged as the best available option. One possible alternative method to AGORA-C was developed and presented in September 2009 is OpenLR. This method aims to become a royalty-free industry standard method for map agnostic location referencing. Although the ROSATTE partners originally did not intend to explore OpenLR as a ROSATTE location referencing method, a number of partners finally did, with positive results. OpenLR location references created from two different maps describing a spatial extend of a speed limit (update), could be decoded on a third independent map with average location matching rates of over 90%. It should be stated that an extensive and detailed testing and validation of this new location referencing technology will be the subject of other projects and initiatives.

4.5. ROSATTE information feedback

- Methods: send back consolidated "data integration report" via email, provide ROSATTE update provider on-line access to their integrated data via a web map interface, set up feedback loop via web service.
- ROSATTE requirements: UR-13, see Annex 1, Table 4.
- ROSATTE selection: set up feedback loop via web service.
- Review: The ROSATTE partners will have access to a web service to collect information upon results of the data integration of the update(s) they have provided.

5. ROSATTE Use Cases/ Reference Implementations

5.1. The Tele Atlas – ROSATTE infrastructure and software tools

The architecture of components developed to test and to validate the ROSATTE data integration is shown in
Figure 6. The overview represents different services (web services for data retrieval, location encoding and decoding, map rendering), processing components, and data files and spatial databases.

The realized implementation can process safety attributes which contain location references in different formats. In addition to standard base64 encoded locations, references encoded in XML (as defined in the AGORA-C specification) and in an XML version defined by French ROSATTE partner for the French BALI and very similar to the AGORA XML, can be decoded. By the end of the project, also an OpenLR location reference decoder was implemented and operational. Different ROSATTE partners provided Tele Atlas with OpenLR location codes. The decoding process matches the map agnostic location references provided by the road authorities to one road element or several connected road elements in the map database. Once this step has been executed successfully, the attributes provided by the road authorities can be added to the map database or they can be used to update existing attributes.

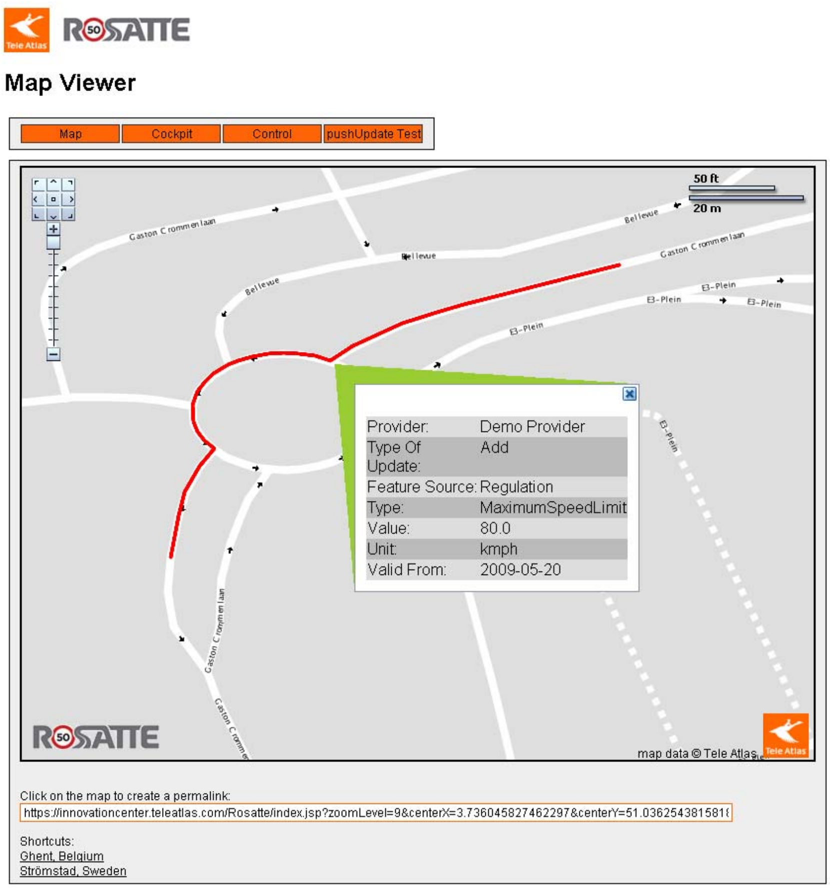


Figure 7 - The on-line map viewer showing an integrated ROSATTE update (maximum speed limit).

In the prototype implementation for ROSATTE data provided by the road authorities and other data providers can be displayed as a map. Since the processed data is stored in a geospatial database, it is possible to easily access and use this data in additional processes in order to integrate them into products which are distributed to the end users of the map data.

5.2. The NAVTEQ Map Reporter based ROSATTE infrastructure

To create the ROSATTE environment, NAVTEQ leverages the system infrastructure of its Map Reporter system, a community portal to interact with the latest map data and to submit reports concerning changes in the real world and deviations between the real world and the road network map database, with the purpose to assist in continuously improving the map database and keeping it up to date.

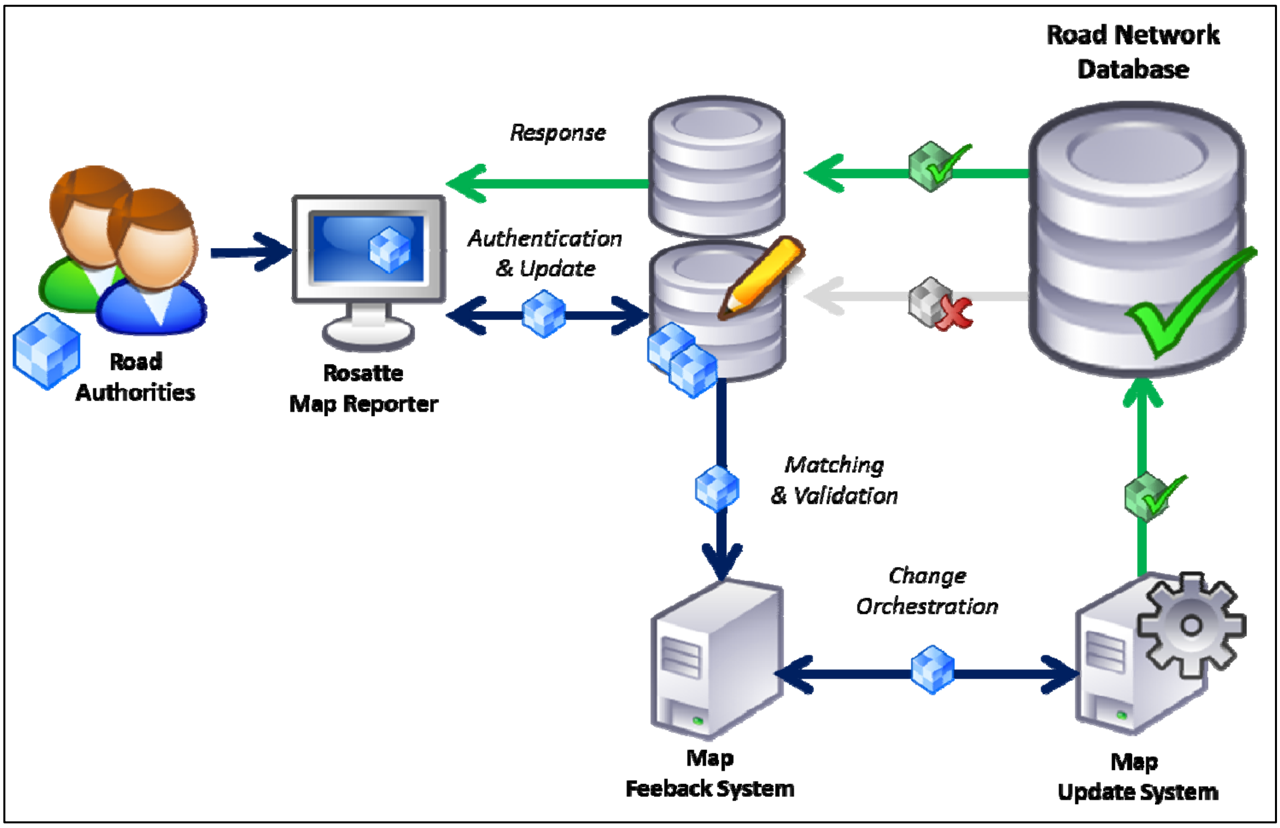


Figure 8 - Architecture of the NAVTEQ ROSATTE infrastructure as part of its Map Reporter community portal.

The Map Reporter Services platform is a complex, interoperable messaging framework for Map Feedback functionality. Using the Map Reporter Services, applications like ROSATTE can publish functions and message to the Map Reporter Backend and participate in the further data integration workflows. The ROSATTE deployment environment is integrated into this live Map Reporter service infrastructure and reuses some key elements of this infrastructure. Its architecture is depicted in Figure 8. The key components are the following:

- Deployment Nodes:
    - Up to 14 virtual and physical machines hosting the following services and public or private endpoints
        - White label ROSATTE Map Reporter viewer and editor
        - SOAP and REST full web service submission and query services

    - Bulk load services
    - AGORA-C encoding/decoding services
    - Middleware API and data queuing services
    - Single sign-on authentication services
    - Online data validation services
    - Feedback tracking and full auditing services
    - Change Orchestration services
    - Reporting and response services
    - Distributed update coordination

- Up to 6 MapTP Smart Map databases for live map editing, distributed layer access and visualization
    - 12 custom layers for ROSATTE specific safety features per road authority

- Synchronized grid of RMOB road network database

Figure 9 provides a screenshot of the NAVTEQ ROSATTE web interface, as component of the Map Reporter environment, which provides functionality for uploading and parsing XML update files, and a map viewer to visualise the results of the data integration.

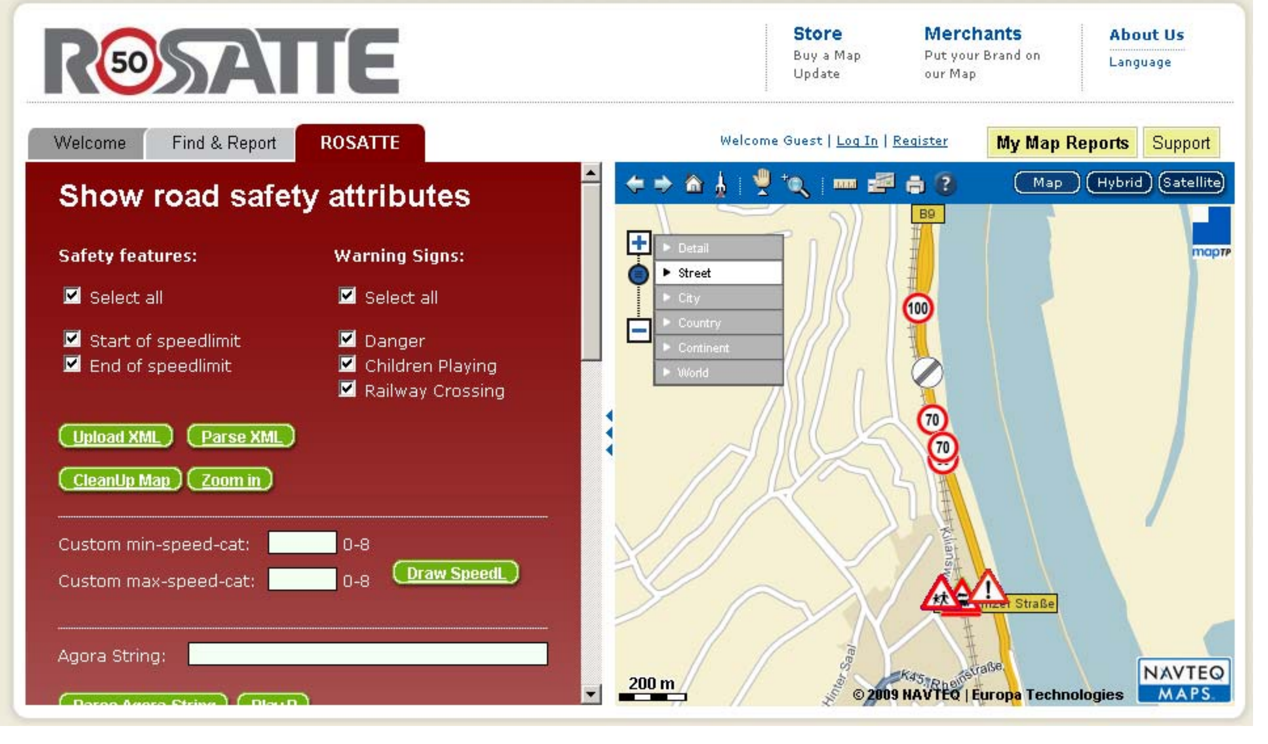


Figure 9 - Web interface for the ROSATTE upload, download and integration service based on the NAVTEQ Map Reporter infrastructure.

An important component of the ROSATTE exchange infrastructure is the AGORA-C location referencing method. The location of each update is encoded as an AGORA-C location reference, which is expressed as a binary base64 encoded string. NAVTEQ provides an efficient web service for encoding and decoding AGORA-C location references. The main service is http link based, and allows automatic encoding of sets of locations from within a

program, and corresponding automatic decoding of sets of AGORA-C locations. In addition, and especially developed for the ROSATTE project, the NAVTEQ AGORA-C web service provides an efficient visual interface for manual encoding and decoding, and for inspection of the results of these processes. Figure 10 provides a screenshot of the NAVTEQ AGORA-C visual interface for manual AGORA-C encoding of locations and decoding of AGORA-C location references.

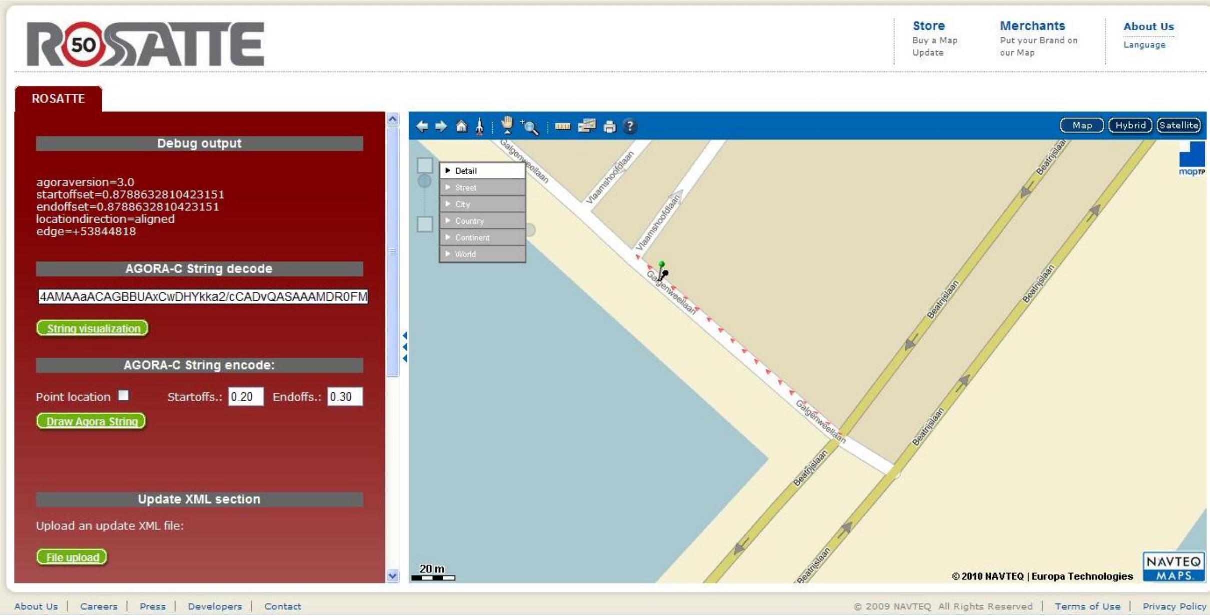


Figure 10 - Web interface for the NAVTEQ service for manual encoding and decoding of AGORA-C location references.

5.3. Data Integration in the Bavarian INTREST operations

In the Bavarian Operation of the 'Intermodal Referencing System for Transport Related Data' (INTREST [14, 15, 16]), the road network itself and certain attributes are updated by data maintenance operations by different INTREST Partners in a decentralised way, Figure 11. These changes need to be integrated with regular updates in the underlying commercial map. This corresponds to a vertical conflation process between an older version of the commercial map, which has been enhanced by user maintenance operations, and a newer version of the commercial map. Data maintenance by INTREST Partners is mostly focused on objects which are not the prime focus of commercial map maintenance (e.g. pedestrian paths and attributes, bike paths, public transport objects, as well as certain POI).

This process allows a 'burden sharing' in data maintenance: certain specific map objects or attributes not covered in commercial digital maps can be maintained regularly, while still benefiting from the overall updating and maintenance of the 'general purpose' digital map.

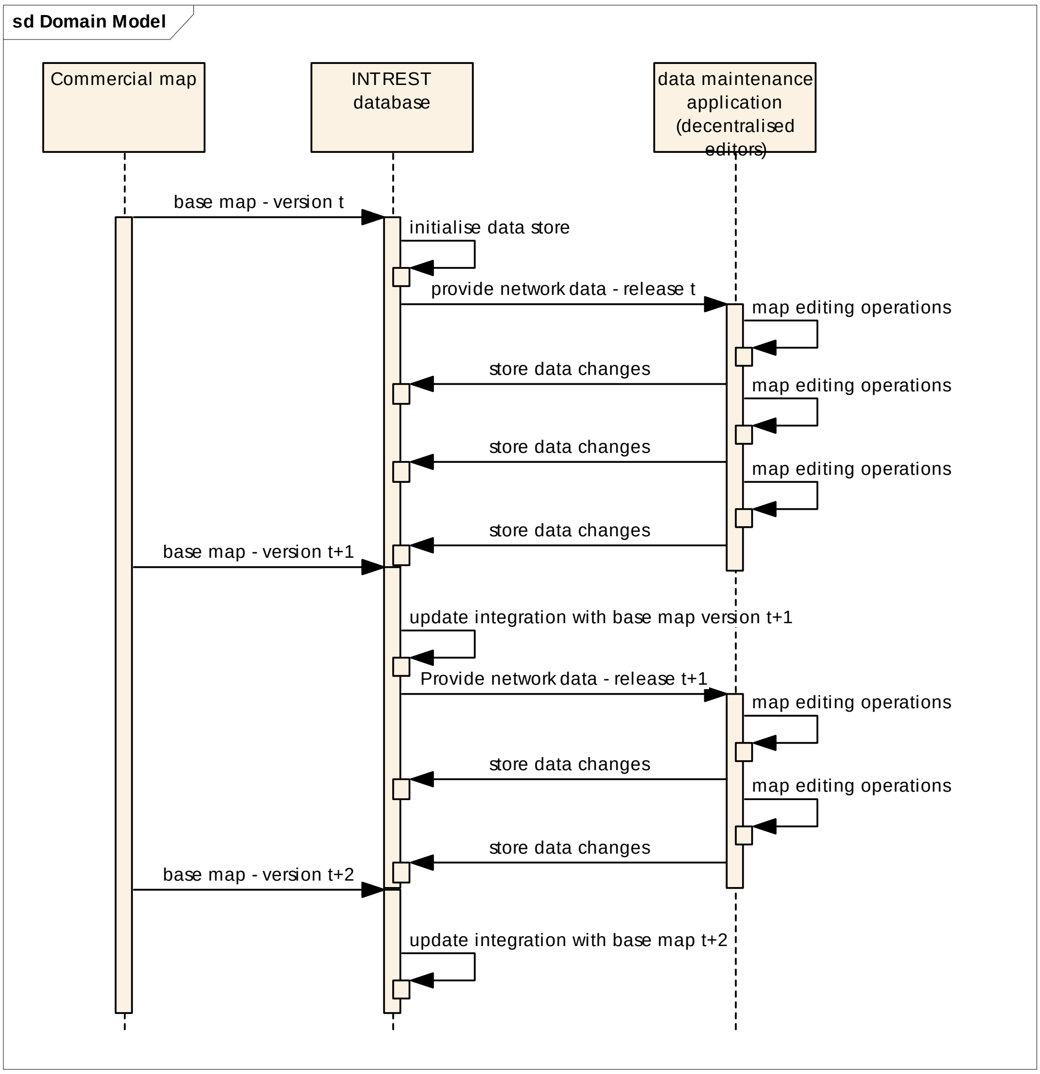


Figure 11 - General process for data maintenance and integration of new base map versions in INTREST.

The conflation process itself consists of a number of steps which have been developed to precisely transfer the updated objects and attributes to the updated base map version. Separate procedures are required for each type of map object. The basic operations for road links include

- ID-comparison: link objects of the base map and objects from data maintenance in INTREST have different ID ranges. With these IDs, new objects from INTREST data maintenance during the previous release period can be identified. This applies to network objects (links, nodes) as well as other map objects (POI, background objects etc.). As a result, new objects from editing are identified.

- Feature Matching: for road links, PTV has developed a matching algorithm which identifies corresponding links in both networks (new version of base map, and version of previous map including data maintenance changes). The Algorithm uses Geometry and link attributes to identify corresponding objects (in the terminology of chapter 2.2. it is hence a feature matching process at object level). Without this matching information, the integration of objects from both sources would lead to double (link) objects in the resulting new INTREST map release.

Both steps allow an identification of differences between the two maps to be merged. From these two steps, a list of insert and delete operations can be created, by which the new base map version can be updated for the changes in the INTREST map version.

- Transfer of attributes: In those areas where specific attributes e.g. of road links are maintained (without ID change) the results of the net-matching procedure are used to identify those objects, for which attributes shall be transferred from the INTREST Map into the new version of the base map.

- Reconstruction of logical references: When all changes from the previous steps are finally operated, all logical references to link objects which are changed during these operations, are reconstructed to point to the new link objects inserted.

The net-matching algorithm plays a particular role in this process.

Basically the net-matching procedure consists of two steps:

- In the first step the deeply digitalized navigation network is reduced to reach approximately the complexity and coverage of the data base, from which additional attributes shall be taken. If the level of detail is similar in the two data bases, this step can be omitted.

- In a second step the corresponding links in both networks are identified (Figure 12). The result is an n:m allocation table containing for each link of the supplier network the corresponding link, or links, of the original network. Identified matches are assigned a matching quality. Since in the INTREST case, both map versions which are conflated come from one provider (either directly or after maintenance manipulations), and automatic matching is therefore not flowed by generalisation or conceptual differences, manual intervention to identify matches can be limited to relatively few objects in the database.

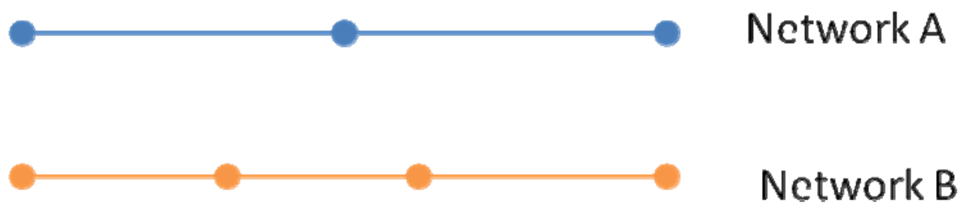


Figure 12 - Corresponding links in the two networks

With regards to ROSATTE operations in Test site Bavaria, the INTREST map which is updated in the process described above, provides the digital network for referencing the safety attributes. With a new release of the INTREST map, ROSATTE safety features need to be re-referenced to the updated network, where changes have occurred. Here again, feature is used to allow afterwards the reconstruction of logical references towards the updated digital road network.

5.4. ASFA Use Case: the Integration of dynamic speed limits.

At the road network operator side, the dynamic speed limits are integrated into the network management system, see Figure 13. Some dynamic speed limits restrictions are integrated into the permanent regulation submitted to validation by the public authority with static speed limits regulation (for tunnel management or traffic control) and other dynamic speed limit restrictions are part of the low (for road works).

Synoptic

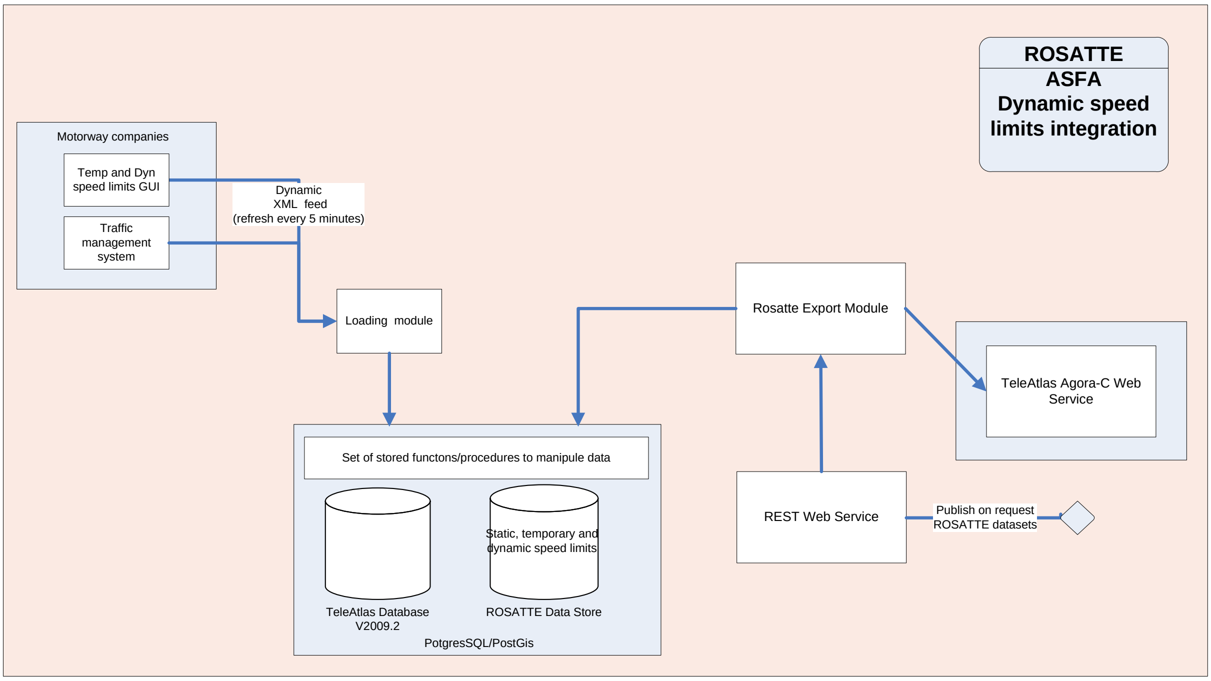


Figure 13 - ASFA dynamic speed limits integration

Input of a speed limit into the system:

Each dynamic speed limit is related to an event as a part of the dataset. Different types of events are considered by the ASFA/AREA test site:

- Works: these events are in place for at least 3 hours. Information about the start time is given to the traffic control centre as soon as the road operator on the field sets up the first beacon. The associated speed limit is automatically calculated, based on configuration of the works site and depending on law (number of lanes left, directions impacted).
- Traffic control in tunnels: these events are in place for 10 minutes minimum. Information about their starting time is given to the traffic control centre as soon as traffic control operator activates traffic control (after receiving an alert for works or incident in the tunnel).

Each event dataset created by the traffic control centre creates an XML dataset.

The speed limits are delivered to the data store operator dynamically through an XML feed.

The feed is automatically refreshed when there is an update (insert, update, delete).

It is refreshed every 5 minutes and sent to the production platform. The content of the feed represents a screenshot of the current temporary and dynamic speed limits on the field. This means that, when a dynamic speed limit is not anymore on the field, it is not anymore in the XML feed. See example of XML feed in Annex 2.

A running process is in charge of:

- loading the XML feed at regular basis,
- updating the ROSATTE data store.

Internal database rules are applied to manage the merge of both dynamic speed limits and existing static speed limits. Major rule is:

<u>"A dynamic speed limit must recover existing speed limits having the same location"</u>

<u>Conversion of locations and extraction of Map IDs</u>:

A speed limit location is defined by a start referent point and an end referent point. In our process they are first converted to geographic coordinates and then mapped on Tele Atlas database to extract map identifiers (MapIds).

The insertion process in the database calls a stored function that will run the conversion.

For example:

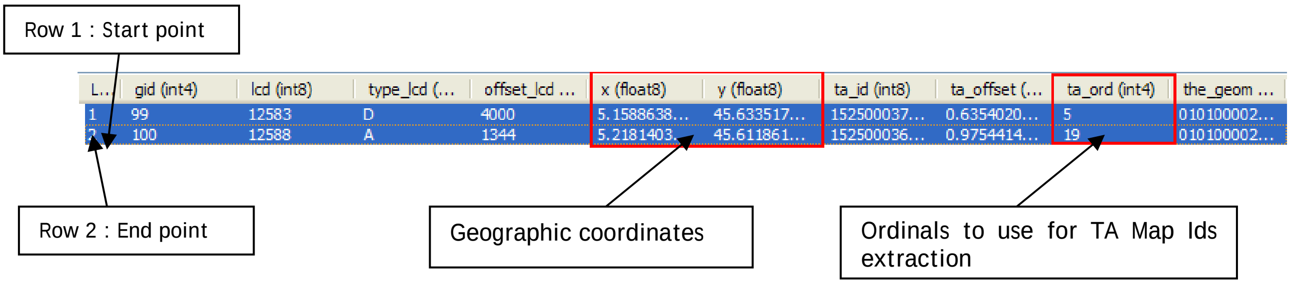
Location: Road = 'A51' Start = km 8 End = km 15 direction = 1

Call to the routine: `at_routepksens_2_xy('A51', 8, 15, 1);`

As a result, two tables are filled with location information:

- *routepksens2xy_result*
- *routepksens2ta_result*

Data in routepksens2xy_result: two rows have been created (the first one for the start point information and the second one for end point) as show in figure 14.

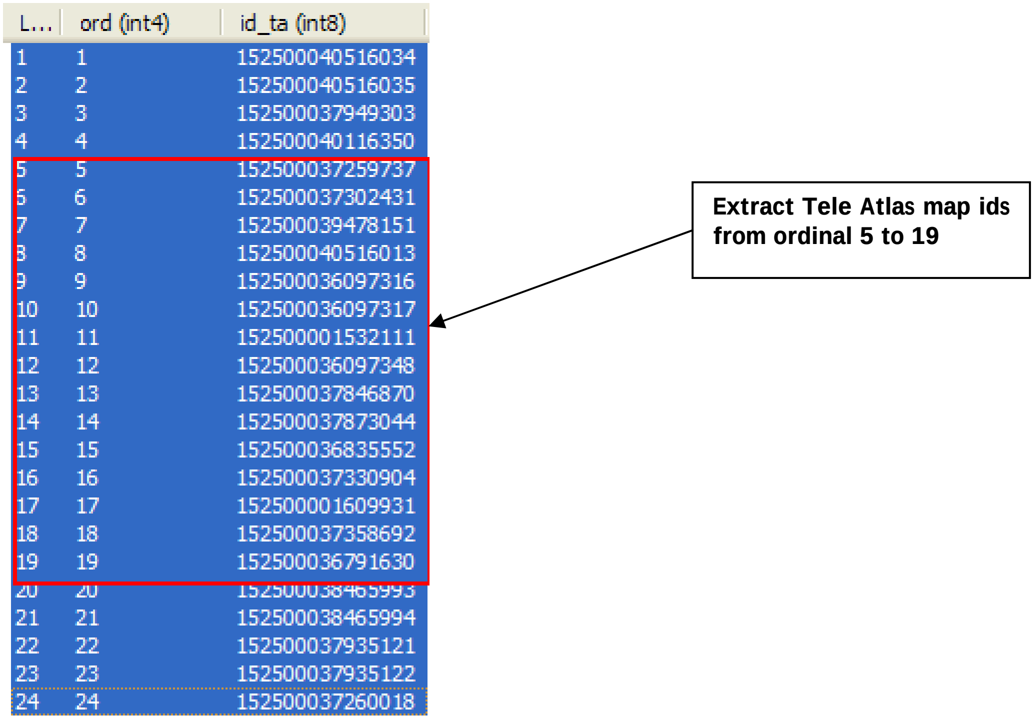


L…	gid (int4)	lcd (int8)	type_lcd (…	offset_lcd …	x (float8)	y (float8)	ta_id (int8)	ta_offset (…	ta_ord (int4)	the_geom …
1	99	12583	D	4000	5.1588638…	45.633517…	152500037…	0.6354020…	5	010100002…
2	100	12588	A	1344	5.2181403…	45.611861…	152500036…	0.9754414…	19	010100002…

Figure 14 - Insertion process

The ASFA extraction process will use the "ta_ord" column value to extract map ids from the second table named *routepksens2ta_result*. In our example, the map Ids should be extracted from ordinal 5 (start of map segment) until ordinal 19 (end of map segment).

Data in *routepksens2ta_result:*



L…	ord (int4)	id_ta (int8)
1	1	152500040516034
2	2	152500040516035
3	3	152500037949303
4	4	152500040116350
5	5	152500037259737
6	6	152500037302431
7	7	152500039478151
8	8	152500040516013
9	9	152500036097316
10	10	152500036097317
11	11	152500001532111
12	12	152500036097348
13	13	152500037846870
14	14	152500037873044
15	15	152500036835552
16	16	152500037330904
17	17	152500001609931
18	18	152500037358692
19	19	152500036791630
20	20	152500038465993
21	21	152500038465994
22	22	152500037935121
23	23	152500037935122
24	24	152500037260018

Figure 15 - Extraction process

The "ord" column value is 5 for start point and Ta_ord value is 19 for end point. Our process extracts list of ids from ord=5 to ord=19, figure 15

Once integrated into the ROSATTE data store, the DSL are processed in the same way than static speed limits.

Export process:

The export process has the following functionalities:

- Extract speed limits information from the ROSATE database, and request for each extracted line, the location information in TA database,

- Send a POST request, for each extracted location in map database, to Tele Atlas Agora-C web service. If the request is successful an Agora-C binary string is returned.

  Example: Location: Road = 'A51'   Start = km 8   End = km 15   direction = 1

  The build request to send to Web Service: see Annex 2.

The AGORA-C binary string received:

ASwBMAAoAiAGBBgXxC4Dipcgfq5cSIMgzUAAAQEDQTQzQAAECQjISAiD/M5TYA==

- Build a ROSATTE dataset in XML format with the extracted information from the DB and the AGORA-C locations (see Annex 2).

Publishing process:

- A REST web service has been created to allow map providers to requests ASFA speed limit datasets. The web service uses the export module to extract data, convert locations in AGORA-C and build a dataset. The dataset is returned as a response to the request.

Feedback process:

- The Web Service can handle feedback that could be sent by data providers. Though it is not clear yet what should be done with this feedback: no further follow up upon received feedback data is foreseen within the scope of the ROSATTE project. The "validation" work package (WP5) will address the result of the integration by comparison of the source data with the integrated data.

6. Conclusions

This deliverable explored viable data integration methods and technologies for map providers to integrate updates provided via the ROSATTE exchange infrastructure. The ROSATTE requirements for data integration presented a guideline during this investigation. The concepts which have emerged during the ROSATTE project, and which have lead to the specification of the ROSATTE exchange infrastructure D3.1, successfully limit the complexity (and cost) and chances to make errors, for map providers to integrate heterogeneous spatial data.

The different use cases (from data providers as ASFA and PTV/INTREST) and reference implementations for data integration (from map makers NAVTEQ and Tele Atlas) presented in chapter 5, which are partly derived from the general data integration framework analysis (chapter 4), proved to be applicable implementations, subject to further testing and validation.

The technical implementations of the data integration components are described in detail in D4.2. These implementations will be tested and validated via a number of test sites, the results will be reported in D5.4.

7. Definitions and acronyms

7.1. Definitions

Term	Definition
(Legal) Traffic regulation	Legal order established by an enacting authority, which regulates the use and equipment of roads e.g. with regards to speed, overtaking ban, traffic lights etc. at a specific location. It often leads to the installation of *traffic signs* at this location. The traffic regulation may cover a point along the road (e.g. pedestrian crossing), a linear location (speed limit along the road from location A to B) or an area location (30km/hour zone, i.e. a set of streets in an area)
Traffic sign	Signs (e.g. speed limit signs) which are put up by road maintenance operators as a manifestation of a traffic regulation for driver information. Traffic signs, by their nature are point objects. To describe a line or an area regulation, several traffic signs may be put up for clarity of information to the drivers. Traffic signs data are often maintained as a part of separate data bases by road maintenance authorities in order to more efficiently manage road (side) equipment. While for certain signs, the content is close in meaning to the corresponding 'safety attribute', transformation is needed to create a safety attribute together with a correct location description. E.g. several traffic signs (repeated speed limit signs) may need to be analysed to define the location/extent of the corresponding safety attribute (speed restriction for several kilometers along a road).
Field survey	Capturing of information by physical road inspection. Measurements/recording of road features including traffic signs etc. on the roads in a specific area or for a specific road level. Field surveys can be done as (repeated) total surveys of areas (full supply) or more ad hoc surveys from small areas. Field surveys may be useful for safety attribute data bases e.g. when creating an initial data set on safety attributes or to support quality mechanisms by providing reference data on safety attributes, which come from other sources.
Road network database	Digital description of road network including certain attributes. There is different ways to describe a road network by data objects. In this document we mean the representation of roads by their geometry (holding a direct location description) and topology using segments and nodes to describe road sections and junctions. In cases where linear referencing methods are used the road network database the data needed for the definition of linear referencing systems may, but is not required to, be part of the road network database. In order to be useful as a basis for on-the fly referencing methods, a road network databases has to include certain structures and attributes (called 'location referencing attributes' in this document) required for all road network databases between which on-the-fly references shall be exchanged.

Term	Definition
Safety Feature / attribute	Feature/attribute in a digital road database which describes the content of a traffic regulation. To be useful, each safety attribute along a road must be paired by the description of its location. The location may be a point, a linear or an area location. To describe the location of a safety features/attributes it can be 'attached' to the road network by (logical) reference to the road database objects in order to clarify their location. Alternatively, a direct location description by coordinates is often used (geo-reference). Its details (as well as the location information) may be directly derived from a traffic regulation (or it could hold a reference to the regulation at its origin). Alternatively, its details (and location information) may be captured by field survey, or from databases including traffic signs. Note: In ROSATTE data stores at enacting authorities or in the digital maps of the information- or map providers, ROSATTE data may be represented either as separate features associated with locations at the road network or as attribution of the road network itself. The term "Safety attribute" in documents D1. and D2.1 refer to both the stored data at either end of the exchange and the exchanged data itself. The term Safety feature in D3.1 refers specifically to the representation of the data which is being exchanged in ROSATTE using the exchange specification. Therefore the terms may therefore be viewed as synonyms since it is the same real world entities that are being represented.

Term	Definition
Location referencing	Location referencing describes a method to describe a location of an object in a digital database. Direct Location referencing uses a description of a location by geo-references (i.e. a description in reference to a geodetic reference system, e.g. latitude and longitude coordinates in WGS84). In the document, geo-references are often used synonymous to direct location references. Indirect location referencing describes a location by its logical reference to other objects (e.g. road segments, or nodes) in a digital database, which themselves hold direct (geo-)references for describing their location. In our case these other objects are those of the road network database (segments, nodes). Indirect references to the road network will also be called 'network reference' in the document. Indirect location references can usually only be interpreted in connection to the specific database objects, which they point to (logical reference). If location references shall be independent of one database, there exist two main ways: ❖ "[It] can be implemented by pre-coding often used locations, as is done in RDS-TMC. The location codes and related additional information are stored in so-called location tables. Advantage is the conciseness of the code; disadvantage the limited amount of addressable locations." (From the AGORA Specification). These predefined and coded locations need to be integrated in each network database, between such location references shall be transferred. ❖ On-the-fly location referencing, which is another method to identify a location independently from a specific instance of a geographic information database. A location known in ones own geographic information set (as indirect location reference) is transformed by an encoding mechanism (using certain rules) into a location code, which can be decoded in reference to another geographic information set (as indirect location reference). The location code contains direct location referencing information (usually geo-references to point locations) as well as typical features/classifications of the (indirectly referenced) data base objects (e.g. classification of a road segment, street names or the like), which are common to all databases. The location referencing method used creates requirements for the source and the target geographic information set with regards to structure and content.
Linear Referencing	The Linear Reference System (LRS, also called Linear Referencing System) is a reference system in which features are localized by a measure along a linear element. Each feature is localized by either a point known as a "milepoint" or a linear event ("segment"). The system is designed so that if a segment of a route is changed only those milepoints on the changed segment need to be updated. (From Wikipedia). In the above classification, LRS is an indirect location referencing technique, which –in the case of a road network – uses 'routes' (a directed chain of segments) as 'aggregated' linear objects, in reference to which a linear location is described either by distance measures (from km 115 to km 120 in reference to the start point at km 0) or percentages (from 55% to 58% of the route length).

Term	Definition
AGORA	is one on-the-fly location referencing method, which is made reference to throughout this document. AGORA requires a digital network description that includes (1) geometric road information, (2) which has a topology to allow routing functions and (3) which contains certain road attributes, such as 'form of way' and 'functional road class'. AGORA was initially developed in an EU funded research project and has evolved into an ISO standard.
Full supply of data	Supply of data with exhaustive comprehensive coverage. A full supply of data is needed when 'initialising' a database, i.e. filling it comprehensively with the data in question. A full supply of data may be established in regular intervals, i.e. asynchronously of data changes. It then replaces the previous data set. The up-to-dateness of the database is only given at moments, where a (new) full supply is available. Road network data are typically updated by full supplies in regular intervals.
Incremental supply of data	Method for regularly updating a database as soon as changes occur in data. The up-to-dateness of the database is continuously maintained. Incremental updates can be provided either through record update events or by comparing different database states. In the case of record update events it is necessary to store information that describes when a safety object is added, modified or deleted. The other option is to send updates as batches of increments created from a process comparing the latest version of a dataset to a previous one.
Conflation	Term used to refer to integration of data by combining the information from at least two spatial data sets to make a new composite data set that is superior to either source data set in either spatial or attribute aspects. In this document the terms integration and conflation are used interchangeably.

7.2. Acronyms

Acronym	Definition
ADAS	Advanced Driver Assistance System(s).
BALI	BAse de données des LImites de vitesse (Speed limit database, French initiative)
Base64	The term Base64 refers to a specific MIME content transfer encoding. It is also used as a generic term for any similar encoding scheme that encodes binary data by treating it numerically and translating it into a base 64 representation. [en.wikipedia.org]
DG INFSO	Directorate-General for Information Society and Media (European Commission).
GIS	Geographic Information System is any system that captures, stores, analyzes, manages, and presents data that are linked to location. [en.wikipedia.org]
LBS	A location-based service is an information and entertainment service, accessible with mobile devices through the mobile network and utilizing the ability to make use of the geographical position of the mobile device. [en.wikipedia.org]
OSF/DCE	Open Software Foundation/Distributed Computing Environment

Acronym	Definition
REST	Representational State Transfer
ROSATTE	Road Safety Attribute Exchange Infrastructure
UML	Unified Modelling Language [www.omg.org/uml]
WP	Work Package
XML	Extensible Markup Language: is a set of rules for encoding documents electronically. It is defined in the XML 1.0 Specification produced by the W3C and several other related specifications; all are fee-free open standards. [en.wikipedia.org]

8. References

[1] ActMAP. Available from: http://www.ertico.com/en/projects/completed_projects/actmap.htm

[2] AGORA-C (2006a). Hendriks, T., Wevers, K., Hessling, M., and H.-W. Pfeiffer, "Specification of the AGORA-C on-the-fly location referencing method", version 1.0, 06 April 2005.

[3] AGORA-C (2006b). Wevers, K., and Hendriks, T., "AGORA-C map-based location referencing", Transportation Research Record, 2006, vol. 1972, pp 115-122.

[4] AGORA-C (2007). ISO, "Intelligent Transport System (ITS) - Location Referencing for Geographic Data-bases - Part 3: Dynamic Location References (Dynamic Profile)", ISO 17572-3, 08-12-2008, International Organization for Standardization (ISO), Geneva.

[5] Beller, A., Doytsher, Y., & Shimbursky, E. (1997). Practical Linear Conflation in an Innovative Software Environment. In: 1997 ACSM/ASPRS Annual Convention and Exposition Technical Papers, Seattle, Washington, April 1997, vol. 2, pp. 146-153.

[6] Blasby, D., M. Davis, D. Kim, and P. Ramsey (2003). GIS conflation using open source tools, white paper.

[7] Chen, C. and C. Knoblock (2008). Conflation of geospatial data, Encyclopedia of GIS, 133-140.

[8] Cobb, M. , M. Chung , H. Foley, F.E. Petry, K. Shaw, and H. V. Miller (1998). A Rule-based Approach for the Conflation of Attributed Vector Data, Geoinformatica, v.2 n.1, p.7-35, March 1998.

[9] Conflex. Available from: www.citygategis.com/conflex.htm

[10] Devogele, T. (1997). Processus d'intégration et d'appariement de Bases de Données Géographiques Application à une base de données routières multi-échelles. PhD thesis, Université de Versailles.

[11] FeedMAP. Available from: http://www.ertico.com/en/projects/completed_projects/feedmap_01.htm

[12] Haunert, J.-H. (2005). Link based Conflation of Geographic Datasets. ICA Workshop on Generalisation and Multiple Representation, 7.-8.July 2005, La Coruña/Spain.

[13] INSPIRE. Available from: inspire.jrc.ec.europa.eu/index.cfm

[14] INTREST. Available from: www.intrest.org.

[15] Landwehr, M., Ortgiese, M. (2003). INTREST – Intermodal referencing system for Transport related data, ERTICO - ITS Europe, technical paper presented at 10th World Congress on ITS, Madrid, 2003.

[16] Thomas Friderich et al. (2004). Data Management System for Transport Planning, Traveller Information and Traffic Control, FOVUS, Aachen, 2004.

[17] JUMP. Available from: www.openjump.org

[18] MapMerger. Available from: www.esea.com/products

[19] MAPS&ADAS, subproject of PReVENT. Available from: http://www.prevent-ip.org/en/prevent_subprojects/horizontal_activities/maps__adas/

[20] OpenLR. Available from: www.openlr.org

[21] REST, Representational state transfer, Available from: http://en.wikipedia.org/wiki/Representational_State_Transfer

[22] ROSATTE, deliverable D1.2 Requirements and Overall Architecture (2008). Available from: http://www.ertico.com/en/activities/safemobility/rosatte/publications/publications.htm

[23] ROSATTE (2007). ROSATTE Consortium, "ROad Safety ATTributes exchange infrastructure in Europe, ROSATTE, Annex 1 - Description of Work, annex to the contract of the ROSATTE Consortium with the European Commission grant agreement 213467), Brussels, 10 December 2007.

[24] Saalfeld, A. (1988). "Conflation: Automated map compilation," Int. J. Geograph. Inf. Syst., vol. 2, no. 3, pp. 217-228, July-Sep. 1988.

[25] SOAP, Simple Object Access Protocol, Available from: http://www.w3.org/TR/soap/

[26] Song, W. T. Haithcoat, and J. Keller (2006). "A snake based approach for TIGER road data conflation," Cartogr. Geogr. Inf. Sci., vol. 33, no. 4, pp. 287-298.

[27] T'Siobbel et al. (2007). Data Sourcing - Advanced Data Collection Technologies. Deliverable D12530, MAPS&ADAS, PReVENT IP – EC - FP6, Available from: www.prevent-ip.org.

[28] Uitermark, H.T. (2001). "Ontology-based geographic data set integration". PhD Thesis. Computer Science Department, University of Twente, Enschede, The Netherlands (http:// www.ub.utwente.nl/ webdocs/ inf/ 1/ t000001e.pdf).

[29] VIVIDSOLUTIONS. Available from: www.vividsolutions.com

[30] Voltz, S. (2006). An Iterative Approach for Matching Multiple Representations of Street Data. In ISPRS Workshop, Multiple representation and interoperability of spatial data, Hanover, Germany, 22-24 February, pp. 101-110.

[31] Walter, V. and D. Fritsch (1999). Matching spatial data sets: A statistical approach, Int. J. Geograph. Inf. Sci., vol. 13, no. 5, pp. 445-473.

[32] Ware J. M., and C. Jones (1998). Matching and aligning features in overlayed coverages, Proceedings of the 6th ACM international symposium on Advances in geographic information systems, p.28-33, November 02-07, Washington, D.C., United States.

[33] Wikipedia. Available from: www.wikipedia.org

[34] Witmer, A. (2000). The Best of Both Worlds: Vector Conflation of Database Segments and Attributes. ESRI User Conference, 2000. Available from: http://proceedings.esri.com/library/userconf/proc00/professional/papers/pap686/p686.htm

[35] Yuan, S. and C. Tao (1999). Development of conflation components, Proceedings of Geoinformatics '99 Conference, Ann Arbor, 19-21 June, 1999.

[36] Zhang, M., Shi, W., Meng, L. (2005). A Generic Matching Algorithm for Line Networks of Different Resolutions. In: Proceedings of the ICA Workshop on Generalisation and Multiple Representation, La Coruña, Spain.

[37] Duckeck, R. et al., 1997. Rules for defining and referencing an Intersection Location (ILOC); Detailed Location Referencing (DLR) for ITS based on ILOCs. ERTICO Committee on Location Referencing, Report, Version 1.0.

[38] Pandazis, J.-C. et al., 1999. Final report. EVIDENCE Project Consortium, Version 2.0.

[39] TPEG. www.tpeg.org

Annex 1: Requirements overview

ID	Requirement name	Short definition	Priority: C: Critical S: Significant I: Of interest	Comments, links to other requirements, open issues
FR-1	Data discovery	A specification of a Discovery service with metadata shall be available.	C	The ROSATTE infrastructure shall provide discovery service with suitable metadata, that enables the Information Provider to easily find services providing road safety attributes. Note: Within ROSATTE WP3, due to time constraints, the discovery service has not been prioritized. The INSPIRE implementing rules for discovery services are believed to be valid also for ROSATTE.
FR-2	Standardized access	Data Services and their use shall be specified.	S	No matter what the content is, accessing and using Data Service's is done the same way across the Europe. Guidelines stating how to access Data Services in a standardized way shall be defined.
FR-3	Data subscription	Guidelines specifying how to subscribe to road safety attributes in the ROSATTE exchange infrastructure shall be provided.	S	Information Providers can subscribe for change notifications for their individual needs. Data subscription functionality is created by combining with FR-5.
FR-4	Specification of Quality management procedures	Guidelines specifying how to quality assure received road safety attributes shall be specified.	C	Guidelines for automatic and semi-automatic quality check routines must be specified. These guidelines should be incorporated into existing procedures.
FR-5	Incremental updates	The ROSATTE infrastructure shall provide both incremental updates and full updates of road safety attributes.	C	Incremental update datasets can be defined using received change notifications.
FR-6	Unambiguous location referencing	The road safety attributes provided through the ROSATTE infrastructure shall be structured to enable unambiguous decoding and interpretation of the referenced locations. Different locating methods allowed.	C	
FR-7	Data Store initiation	The project shall provide guidelines for Data Store design and initiation.	I	
FR-8	Data import	The project shall define guidelines for import of road safety attributes and road network data. If suitable import tools are non-existent, new tools shall be developed.	S	

ID	Requirement name	Short definition	Priority: C: Critical S: Significant I: Of interest	Comments, links to other requirements, open issues
FR-9	Workflow support	The project shall produce a specification of tools and guidelines for integrating data maintenance with legal workflow.	I	
FR-10	Presentation and maintenance tools	The project shall develop specifications of how to present and maintain the road safety attributes. If existing tools are not suitable, new tools shall be developed.	C	Where suitable existing tools do not exist, tools that enable presentation, maintenance and publishing of the road safety attributes must be developed and implemented. The tools shall be built on existing work and standards where suitable.
FR-11	Feedback loop	A feedback channel from information providers back to enacting authorities shall be provided.	C	
FR-12	Integration tools	Tools to integrate road safety attributes into existing information providers systems shall be developed if existing tools does not provide the satisfactory functionality.	S	Generic software components must be developed to integrate the road safety attributes in a quality assured (ref FR-4) and automated manner.
FR-13	Flexible type definitions	The meaning of “Road safety attributes” is not finally decided. Changes will occur in the future.	C	It shall be possible to add and change (to some degree) the available type definitions describing road safety attributes.
NFR-1	Availability	Valid quality parameters related to availability shall be declared in the metadata associated with the road safety attributes.	C	Degree to which geographic data is available at a certain place and at a defined time. Possible quality parameters: - Communication failure rate These quality parameters should be specified in the metadata attached to delivering system.
NFR-2	Up-to-dateness	Valid quality parameters related to up-to-dateness shall be declared in the metadata associated with the road safety attributes.	C	Degree of adherence of geographic data to the reality changing with time. Possible quality parameters: - Date of last update - Date of origin - Rate of change
NFR-3	Completeness	Valid quality parameters related to completeness shall be declared in the metadata associated with the road safety attributes.	C	Degree of availability of all information needed to describe the reality. Possible quality parameters: - Missing data - Surplus data

ID	Requirement name	Short definition	Priority: C: Critical S: Significant I: Of interest	Comments, links to other requirements, open issues
NFR-4	Correctness	Valid quality parameters related to correctness shall be declared in the metadata associated with the road safety attributes.	C	Degree of accordance of geographic data (feature(s), attributes, functions, relationships) to corresponding elements in reality, up-to-dateness being presumed. Possible quality parameters: - Geometric correctness - Topological correctness - Thematic correctness
NFR-5	Consistency	Valid quality parameters related to consistency shall be declared in the metadata associated with the road safety attributes.	C	Degree of accordance of geographic data (data structure, their features, attributes and relationships) to the models and schemas (conceptual model, conceptual schema, application schema and data model). - Geometric consistency - Topological consistency - Thematic consistency
NFR-6	Accuracy	Valid quality parameters related to accuracy shall be declared in the metadata associated with the road safety attributes.	C	Degree of adherence of geographic data to the most plausible or respectively the true value. - Absolute position accuracy - Relative position accuracy - Quantitative attribute accuracy
NFR-7	Reduced data update delay	The time delay from the moment a Public Authority regulation is effective, until the end user data have been updated, shall be reduced.	S	Related to NFR-2. The infrastructure itself may have minor delays, but the administrative routines on public authority side must be adapted to the lifetime of the data handled. Update frequencies of 24 hours or less is a reasonable requirement.

ID	Requirement name	Short definition	Priority: C: Critical S: Significant I: Of interest	Comments, links to other requirements, open issues
CR-1	Conformance with European law.	The ROSATTE infrastructure shall offer its services in a way that conforms with the INSPIRE directive. This includes creation and maintenance of metadata, a discovery service using it with a minimum set of search criterions, view services, download services and supporting services.	C	SOA Web Services Minimum metadata elements required: Identification (Name,type,URL) Classification Keyword Geographic location Temporal reference Quality and validity Conformity. Access conditions Access limitations Responsible organization Minimum search criteria: Classification Keywords Geographical location Quality and validity Access conditions Responsible organization **Links to functional requirement FR-1**

Table 2 - List of System requirements in ROSATTE

User requirement		Additional explanation	Roles involved
Number/ID	(short) Definition		
UR-1	Data discovery	Map providers need to be able to find providers of available road safety attributes.	Information Provider
UR-2	Data subscription	Map providers need notifications when relevant data changes.	Information Provider
UR-3	Unified access	Map providers need a unified way of access and retrieval of road safety attribute data across the Europe.	Information Provider
UR-4	Data updates	Map providers need both full and incremental updates of road network safety attributes, expressed according to one unified data model. Updates can be initiated by change notifications from road authorities.	Information Provider
UR-5	Location referencing	Map providers need a location reference which enables unambiguous decoding and interpretation of the referenced location.	Information Provider
UR-6	Quality	Map providers need quality assured data to integrate into their own databases in order to ensure the quality of the end user products	Information Provider
UR-7	Update delays	Map providers need notifications and data updates at a rate that is suitable compared to the lifetime of the affected data.	Information Provider

Table 3 - Map Provider requirements

User requirement		Additional explanation	Roles involved	Partner use cases
Number/ID	(short) Definition			
UR-7	Data Store initiation	If missing, road authorities need guidelines for data store design and initiation. (These guidelines are one expected result of WP 2.)	Data Store Operator	LRAd1
UR-8	Initial supply / Data import	Road authorities need a way to import road network and road safety attributes from different sources. This includes both the initial supply and updates.	Data Store Operator	AS5, LRAd2-LRAd6, LRUp9
UR-9	Integration of the attribute supply in to the work flow of regulations	Road authorities need a way to integrate data maintenance into the legal work flow, with minimum extra effort.	Data Store Operator	AS1, AS2, AS3, FL0-FL11, LRUp9, Np09-Np11, Np19-Np21, Ob1, Ob2, Sr1-Sr10
UR-10	Data presentation and maintenance	Road authorities need tools for data presentation and maintenance.	Enacting Authority	AS4-AS6, Sr1, Sr3, Sr8 Ob2, Np17, LRUp7, LRUp8, Np01-Np08
UR-11	Data publishing,	Road authorities need a data	Enacting	AS7, Np12,Ob4,

User requirement		Additional explanation	Roles involved	Partner use cases
Number/ID	(short) Definition			
	both for the ROSATTE infrastructure, and for public websites.	publishing mechanism which is flexible and easy to adapt.	Authority	Sr2
UR-12	User feedback	Road authorities need feedback from users to improve quality.	Information Provider, Enacting Authority	AS8, FL10, Np18, Ob6
UR-13	INSPIRE conformance	INSPIRE has become a directive, and conformance is a requirement for data owners.	Enacting Authority	
UR-14	Existing work and standards	To protect investments, ensure acceptance, and save time and effort, work should build on existing work and standards.	Enacting Authority, Data Store Operator	Derived from Technical Annex
UR-15	Quality management	Road authorities need quality management in order to guarantee the provision of quality assured data.	Enacting Authority, Data Store Operator	LRUp10, Np16, Ob5, Sr10

Table 4 - Road authority requirements

9. Annex 2: ASFA XMLs

Example XML feed

```
< VITESSE_TMP>

        <!- nom de l'exploitant et du concessionnaire -->
        <CONCESSIONAIRE>AREA</CONCESSIONAIRE>
        <EXPLOITANT>AREA</EXPLOITANT>

        <AXE nom= 'A43' sens='2'>
                &lt;SEGMENT'&gt;
                        <!- début, fin du segment-->
                        <DEBUT pk='50.6' x='86.145' y='20.457'/>
                        <FIN       pk='125.995' x='64.408' y='56.6547'/>
                        <!- vitesse normal (statique) -->
                        <VITESSE_NORMAL>130</VITESSE_NORMAL>
                        <!- vitesse temporaire-->
                        <VITESSE_LIMITE>90</VITESSE_LIMITE>
                        <!- cause en code phrase datex -->
                        <CAUSE>RWK</CAUSE>
                        <!-- appliquée aux types de véhicule … -->
                        < !--'o' pour oui, 'n' pour non →
                        <VEHICULE vl= 'o' camion='o' autre='n'/>

                &lt;/SEGMENT&gt;

                &lt;SEGMENT&gt;
                        <DEBUT pk='125.995' x='64.408' y='56.6547'/>
                        <FIN        pk='190.321' x='13.596' y='20.779'/>
                        <!- vitesse normal (statique) -->
                        <VITESSE_NORMAL>130</VITESSE_NORMAL>
                        <!- vitesse temporaire -->
                        <VITESSE_LIMITE>110</VITESSE_LIMITE>
                        <CAUSE>RWK</CAUSE>
                        < !--'o' pour oui, 'n' pour non →
                        <VEHICULE vl= 'o' camion='n' autre='n'/>

                &lt;/SEGMENT&gt;
                ...
        </AXE>

        < !—eventuellement autres axes / segments ici →

</ VITESSE_TMP>
```

The built request to send to Web Service:

```
<?xml version='1.0' encoding='UTF-8'?>
<LocationTranslationRequest xmlns='http://www.location.com/locationtranslation'
xmlns:adtlt='http://www.location.com/adtlt' version='1.0.0'>
        <LocationTranslationQuery vendor='TELE ATLAS'>
                <adtlt:Location release='2009.02'>
                        <adtlt:MapIdString delimiter=',' >
                        152500037259737,152500037302431,152500039478151,152500040516013,1525000
                        36097316,152500036097317,152500001532111,152500036097348,15250003784687
```

```
                    0,152500037873044,152500036835552,152500037330904,152500001609931,15250
                    0037358692,152500036791630</adtlt:MapIdString>
            </adtlt:Location>
            <TranslationOutput format='AGORA-C' version='3.0'/>
        </LocationTranslationQuery>
</LocationTranslationRequest>
```


ROSATTE dataset in XML format containing extracted information from the speed limit database and the AGORA-C locations.


```
<?xml version="1.0" encoding="iso-8859-1" ?>
<rst:ROSATTESafetyFeatureDataset xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:gml="http://www.opengis.net/gml" xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:TPEG="TPEG"
xsi:schemaLocation="http://www.ertico.com/en/subprojects/rosatte/rst/Rosatte.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" gml:id="514d7f99-108a-4328-a0f2-
94ed70caba20" xmlns:rst="http://www.ertico.com/en/subprojects/rosatte/rst">
 <gml:featureMember>
        <rst:GenericSafetyFeature gml:id="fcd81528-775b-4642-bc49-6c20bcfbe05b">
        <rst:id>
              <rst:SafetyFeatureId>
                       <rst:providerId>ASFA-AREA</rst:providerId>
                       <rst:id>100303</rst:id>
                </rst:SafetyFeatureId>
         </rst:id>
        <rst:locationReference>
               <rst:AgoraLocationString gml:id="dae96a82-01fe-4dec-8854-81a815958c75">
          <rst:base64String>ASwBMAAoAiAGBBgXxC4Dipcgfq5cSIMgzUAAAQEDQTQzQAAECQjISAiD/M5TYA==<
          /rst:base64String>
              <rst:agoraBinaryVersion>3.0</rst:agoraBinaryVersion>
              </rst:AgoraLocationString>
         </rst:locationReference>
         <rst:validFrom>2010-01-21</rst:validFrom>
        <rst:UpdateInfo>
              <rst:type>Add</rst:type>
         </rst:UpdateInfo>
         <rst:source>Regulation</rst:source>
        <rst:encodedGeometry>
               <gml:Point gml:id="18d0bd06-5fb6-495f-96d4-3f9c680c3b18" srsName="start">
                <gml:coordinates> 5.1588638188954 45.6335170765245</gml:coordinates>
                </gml:Point>
               <gml:Point gml:id="12e44314-c869-433a-ad14-302e5cd01bfc" srsName="end">
                <gml:coordinates> 5.21814033505741 45.6118616401037</gml:coordinates>
                </gml:Point>
         </rst:encodedGeometry>
         <rst:type>SpeedLimit</rst:type>
        <rst:properties>
              <rst:SafetyFeaturePropertyValue>
                    <rst:type>MaximumSpeedLimit</rst:type>
                    <rst:propertyValue>
                           <gml:measure uom="kmph">130</gml:measure>
                    </rst:propertyValue>
              </rst:SafetyFeaturePropertyValue>
         </rst:properties>
         </rst:GenericSafetyFeature>
         </gml:featureMember>
</rst:ROSATTESafetyFeatureDataset>
```