Final report

"Methodology for establishing general databases on transport flows and transport infrastructure networks"

MESUDEMO
Contract No: ST-97-SC.1186

Project Co-ordinator:

Agder Research Foundation, AF

Partners:
NEA Transport research and training, NEA
National Technical University of Athens, NTUA
Istituto Nazionale Di Statistica - Italy, ISTAT

Date: August 2001

PROJECT FUNDED BY THE EUROPEAN COMMISSION UNDER THE TRANSPORT RTD PROGRAMME OF THE 4TH FRAMEWORK PROGRAMME

PREFACE

This Deliverable summarises the main findings in the MESUDEMO project. Totally ten Deliverables have been produced in MESUDEMO.

The final report has been produced as a joint effort by all the Consortium partners; AGDER Research Foundation (Norway), Istituto Nazionale Di Statistica (Italy), National Technical University of Athens, Department of Electrical and Computer Engineering (Greece), NEA Transport Research and Training (the Netherlands) and with considerable contributions from representatives of the Ministry of Transport of the Netherlands, of SES in France and FIT Consulting in Italy.

This report is edited by AGDER based on written contribution from

AGDER: Thorkel Askildsen, Dag Bjørnland, Jørn Cruickshank, Anders Langeland, Gunnar Stavrum
NEA: Philippe Tardieu, Adrian Vilcan
NTUA/DECE: Elias Koukoutsis
AVV: Jos Bekkers, Peter M. Mak
FIT Consulting: Massimo Marciani

Kristiansand 28 August, 2001

CONTENTS:

THE PARTNERSHIP

At the outset in September 1997 the MESUDEMO Consortium consisted of five partners:

- ✓ AGDER Research Foundation (Norway) (co-ordinator)
- ✓ CBS - Copenhagen Business School (Denmark)
- ✓ ISTAT - Statistical office of Italy
- ✓ NEA Transport Research and Training (the Netherlands)
- ✓ NTUA/DECE - National Technical University of Athens, Depart. Of Electrical and Computer Engineering (Greece) – for simplicity called hereon NTUA.

Due to staff changes CBS decided to withdraw from the project, and terminated its project activities by the end of 1998. NEA continued the work of CBS together with AVV. ISTAT sub-contracted FIT Consulting to carry out major parts of the work on their behalf.

In early spring 1999 the Commission (DG VII) decided to follow the advice of the Consortium and include a Transalpine Pilot Study as a new work package (WP 10) within the logical frame of MESUDEMO. This modification of the original work programme had consequences for the originally planned content of WPs 4, 8 and 9 and necessitated input from CONCERTO (NEA) and ASSEMBLING (MCRIT).

Participants with considerable contributions given during the project:

- ✓ AVV, (Part of) Ministry of Transport, the Netherlands
- ✓ (Part of) Ministry of Transport, Austria, related to WP 10
- ✓ SES, (Part of) Ministry of Transport, France.

During the scientific work period, September 1997 - August 2000, the partners had 18 Consortium meetings discussing drafts etc. All partners contributed to the work of the various WPs, and have consequently been involved in all the Deliverables. WP leaders with specific responsibility were appointed for the various WPs:

WP1: ISTAT
WP2: NEA
WP3: AGDER
WP4: ISTAT (NTUA for computer issues)
WP5: AGDER
WP6: NEA (originally CBS)
WP7: NEA (originally CBS)
WP8: NTUA
WP9: AGDER
WP10: NEA (NTUA for data architecture)

The Consortium has produced the following Deliverables and reports:

Del. 1 + 5: *Analysis of existing methods, flows and infrastructure databases*
Del. 2 + 3: *Identifying data on goods and passenger flows*
Del. 6 + 7: *Definition of networks to be reported and selection of network for the database*
Del. 4 + 8: *Data architecture and strategic management. Computer models and data architecture*
Del. 9: *Ensuring compatibility between supply and demand side methodologies*
Del. 10: *Transalpine Pilot Study and demonstrator.*

Report, Workshop 1, May 1998, Copenhagen
Report, Workshop 2, June 1999, Rotterdam

EXECUTIVE SUMMARY

Introduction and overview

The MESUDEMO project deals with tasks 4 and 5 in the Strategic Transport Research Section of the 4th Framework Programme in the European Union (EU). As a RTD (Research and Technical Development) project it is concerned with methods for compiling, storing, maintaining and disseminating information on transport networks and on flows of goods and of passengers. Developing data architecture and data models adapted to these flows and networks has been a major challenge. In the chain of events comprising causes of transport, transport performance and the impacts from transport, the emphasis in MESUDEMO is on transport performance. Such boundaries between cause and effect are more or less artificial, but might be a necessary means preventing projects from becoming unduly large and unmanageable.

The INFOSTAT project laid the foundation for a European Transport Policy Information System known as ETIS. In the MESUDEMO project those parts of ETIS called transport demand indicators and transport network indicators have been taken as a starting point.

Transport statistics provides useful information both on the process of monitoring transport development, but it is also useful for transport policy decision support. In order to extract policy relevant information from existing primary data, it is necessary to select from and combine collected data into indicators and variables that can represent the phenomena under investigation. The terms *indicator* and *variable* were used rather interchangeably in the INFOSTAT project. The ranking does not matter in MESUDEMO and, therefore, no strict ordering has been observed. Mostly the term indicator is used in MESUDEMO.

The ensuing figure introduces the various work packages (WPs) of MESUDEMO and their interlinkage. Ensuring compatibility between transport flows and networks and between these indicators and data models and data architecture is crucial to the success of MESUDEMO. Therefore, ensuring compatibility has been identified as a separate work package. Separate identification has also been given to the important transport pilot study called the Transalpine Pilot Study.

As indicated in the figure, MESUDEMO was originally designed mostly as parallel strings of work packages. WPs 1-4 represent one string on the demand side of transport, and WPs 5-8 cover the supply side. As often experienced, however, knowledge is not generated in such a linear fashion. On the contrary, experiences from later work packages impact on preceding ones, spurring further work there. A non-linear and iterative form of knowledge generation has been pronounced in the case of MESUDEMO. This fact may to a certain extent explain why all deliverables have been published in the last few months of the project work, although scientific work commenced in the autumn of 1997 and was finished by August 2000.

All in all ten deliverables have been produced in the project, summarising the results achieved in the ten work packages, in addition to two reports from workshops held in the course of the project.

A project life of nearly three years is a very long time span in applied research and puts particular strain on a research consortium organised for the work. MESUDEMO has experienced this in full, for instance with one consortium partner withdrawing completely from the project. From time to time participating organisations have been in need of putting new team members on the project.

The Work Packages and organisation of MESUDEMO

The policy context of MESUDEMO

The Common Transport Policy (CTP) of the EU has been fundamental in the project work in more than one way. There is a need for strategic transport research in order to provide the necessary knowledge to support policy actions in the short-term as well as for long-term strategic policy actions aiming at structural changes.

Current problems in the transport sector can be classified in the following manner:

- Economic inefficiency, manifesting itself as increasing transport costs and inadequacy of transport networks to cope with growing logistics requirements and mobility needs of transport users.
- Congestion problems, particularly in central parts of Europe and in larger metropolitan areas.
- Intolerably large numbers of traffic accidents leading to much suffering and implying high private and social costs.
- Environmental damage and consumption of non-renewable resources.
- Spatial and regional imbalances.

Developing and diffusing a unified European transport policy that appears credible to all European countries, is an extremely important and difficult task as scale and scope of mobility patterns have evolved out of different social and economic histories. Still European countries become economically, socially, and politically more integrated and in need of a common transport policy. At the same time as social and economic processes are becoming functionally more integrated than before, they are played out over an extended geographical arena: Local problems do not necessarily have local causes, general solutions to specific problems may create adverse effects in different localities and so forth.

An important component of the CTP is to facilitate interoperability and intermodality in transport, thereby achieving more efficient, cost-effective and environmentally friendly use of the networks in the different modes of transport. Such a development will promote sustainable mobility that is fundamental to the CTP. However, no current transport information system, or any information system representing social systems for that matter, provides such data to any noticeable extent. The challenge to projects such as MESUDEMO is for that reason formidable.

Transport research must respond to policy priorities, but as the political perception of transport also is coloured by deliveries from transport research, this is not just a one-way relationship. A gradually evolving refinement of the perception of the way the transport system develops, hinges on interrelationships among decision-makers, planners, researchers and other actors within the system. This has been clear in the minds of the partners of the MESUDEMO consortium.

The aim of MESUDEMO has been to develop tools that will cater for decision support at the European transport policy level. To attain this there is a need to harmonise transport information generation across Europe and to establish data architecture and data models for the production of relevant information. In MESUDEMO the conclusion has been drawn that information can best be supplied at the national level, but processed at different levels.

As the scope of problems as well as that of priorities varies from the local to the global level and as data are mostly compiled at the national level, the recommendations from MESUDEMO may be viewed as a combination of a top-down design approach and a bottom-up design process.

Top-down part of the ETIS design:

Policy issues → Formulation of the questions to which ETIS has to provide answers → Initial decisions on the kind of desired data and algorithms that have to be included in ETIS → Compilation of needed data and construction of data architecture → Giving answers to policy questions.

Bottom-up part of the ETIS design:

Determination and harmonisation of available data or data that can be collected → Determination of the possible algorithms and processing methods that can be successfully used → Identification of the questions, to which the resulting information system can answer.

Therefore, the actual procedure of creating any ETIS comprises two stages. There is the problem-driven, top-down initial one that has to be complemented by a data-driven, bottom-up one, which will actually determine the parts of the top-down procedure which can actually be implemented at any point in time. The two stages, therefore, result in the actual ETIS, that materialises at any point in time.

Policy priorities, planning issues and information technology change as time goes by. National backing of ETIS is essential, as data will mostly be compiled at that level. ETIS must deal with non-compatible data, both technically and conceptually, since collection of existing primary data is motivated by various policy challenges between countries and between modes. ETIS also is faced with the challenge of adjusting to future changes both with respect to mobility and the development of transport performance which influences the policy challenges in the future. Therefore, the MESUDEMO consortium has come to the conclusion that ETIS has to be both simple and flexible to work with, simple to maintain and easy to revise.

Project strategy development (WP 1 and 5) and research methods

Desk research and literature studies initiated the project work and were extensively used all through the project. Initially desk research and literature studies were supplemented with a questionnaire survey on current database practice. A sample of 99 respondents was drawn from all over Europe, and the survey had a reasonable response rate. The questionnaire survey contributed to the MESUDEMO project in several ways:

- It was a useful process to raise awareness of the whole project; thus the construction of the questionnaire represented a starting point for the work on all work packages.

- The results from the survey constituted a basis for the description of existing methods and an introduction to the average practice in the compilation and the use of transport information systems in Europe.
- The general impression from the responses was a necessary departure point for selecting interview objects for in-depth interviews later on in the process and for establishing contacts with transport database holders.
- The survey provided the project with inputs about possible solutions for a Pan-European database on transport infrastructure and transport flows.

The questionnaire survey lead to thirteen in-depth interviews, including international organisations in the quest for best practice. In the in-depth interviews it was possible to get a grasp on the functionality and organisation of complete systems. This was an important input in the project work because it gave insight into current methods for the creation of ETIS and into what ETIS might look like in practice.

Pilot studies were considered essential to the success of MESUDEMO and served several purposes. It was a way of testing hypotheses developed in the project. The questionnaire survey and the in-depth interviews could not go deep enough as to the feasibility of the ideas of the architecture of ETIS. Then there was the issue whether the concept of one standardised, integrated, Pan-European system was possible, and if it could be financed and maintained.

Pilot studies could produce knowledge about what could be done with existing data and data models, and whether it would be worthwhile spending resources on making one Pan-European transport policy information system based on existing data and tools.

Two pilot studies and one study on data availability were carried out. A pilot study called TRANSITIE was carried out within the work package on goods flows (WP 2) and a study on passenger transport flows within the work package on passenger flows (WP 3). The most ambitious pilot was the Transalpine Pilot Study (WP 10) which also included a demonstrator showing how existing databases and transport model elements could be combined to give answers to given policy issues in the Alpine region, presented in a Geographical Information System (GIS).

Workshops were considered useful for co-ordinating project work as well as generating ideas. They were considered both to be instrumental for scientific progress within on-going work packages, as well as a first introduction to ensuing work packages. They also functioned as an instrument for disseminating acquired knowledge between each WP, as well as between the MESUDEMO project and related RTD projects and institutions. Two workshops were held during the project work, one in May 1998 and the other in June 1999.

Transport flows (WP 2 and 3)

The seven goods flow indicators and the eleven passengers flow indicators from the INFOSTAT project were analysed with the objective of making them operational in a three stage research process.

In the first stage the indicators were analysed conceptually and an operational level selected.

Transport is carried out in space and therefore, all flow indicators must have a spatial dimension. In MESUDEMO the indicators can include an origin and a destination (O/D) designation or either an origin (O) or a destination (D) designation. For that purpose an administrative classification system of space called the NUTS system (Nomenclature des Unités Terrioriales Statistiques), developed by the EU Commission, has been used in MESUDEMO to illustrate the spatial reasoning.

The NUTS system is specified for several spatial magnitudes. In NUTS 0 the zone is equal to the territory of the national state. So an O/D specified for NUTS 0 means that the whole of one country is considered one area. Zones with higher numbers represent smaller areas within each country. For each country the NUTS level and zone boundaries have been defined and are known.

When ETIS has been fully developed at some NUTS level, it will be possible to aggregate up NUTS levels according to the problem to be solved or the question to be answered. Specification for more details might be impossible when the NUTS level of ETIS once has been set.

NUTS 3 is a zone size at which many of the policy questions asked, can be answered at a sufficient level of accuracy and detail. NUTS 3 is also the level at which several nations have chosen to develop their national transport planning models. In practice NUTS 2 might be the best to aspire for, given the present state of affairs in databases.

In particular for the goods flow indicators it was considered essential to implement a transport chain concept. Then simply an O/D designation does not suffice, but one or more transshipment points must be identified to cope with intermodality and multimodality issues.

In stage two a comprehensive review procedure was used to identify the level of compatibility between the actual proposed indicators and the structure of data available from different sources. The conclusion from the extensive search for available and relevant statistical information was rather simple:

As far as goods are concerned, traditional statistical data on European goods transport flows are mostly restricted to a unimodal registration at an annual basis of tonnes and/or tonne-kilometres broken down by commodity group and only specifying country of origin and country of destination (NUTS 0). For passenger transport the conclusion was even simpler. Available information on passenger transport demand on the international level is poor, apart from air and maritime transport. There are very few O/D related data available for border crossing transport. Some national travel surveys specify O/D by transport modes on a NUTS 0 level for neighbouring countries, i.e. countries with substantial contact between one another.

The lack of more detailed O/D information is seen as one of the major shortcomings of existing transport statistics.

The third research stage in the context of work package 2 and 3 comprised two studies on the perspectives of harmonising and merging data from various sources. In the case of goods transport the pilot study was called TRANSITIE and used trade and transport data from the Netherlands in an effort to create a transport chain database. Such a linking of different databases for all transport modes was outside the scope of MESUDEMO, and maritime transport was selected as an interesting mode in the Dutch case.

Several experiments were carried out. The one which yielded the best result, took into account:

- direction of trade flows (incoming/unloading or outgoing/loading in the Netherlands)
- country of origin/loading or destination/unloading (approximately 200 countries)
- Dutch region of origin/loading or destination/unloading (54 regions that can be aggregated to NUTS 3).

This alternative resulted in the linking of 81 per cent of the flows, a proportion applying to exports as well as to imports. It seemed that an initial target set at 90 per cent could be reached if further attention would be given to the Dutch foreign trade.

Considering the difficulties experienced in the Netherlands, the creation of a Pan-European consistent transport chain database will be rather challenging and complicated. It cannot be ruled out that similar data problems also exist in other European countries. However, by comparing data from different sources it is at least possible to check the quality of data. A quality check on the transport/trade data is recommended for all countries.

The study on the availability of passenger flow data showed that the existing situation would make the road towards ETIS complicated, but not impossible. Some countries perform national travel surveys and collect/model passenger flows from available data. Some countries also compile statistical information on international trips.

Existing national data are currently too heterogeneous to be harmonised and too scarce to provide a satisfactory European coverage within ETIS. A great effort is consequently needed to get the different countries to enrich and adapt their present systems of passenger data collection. By combining both a top-down and a bottom-up approach, and gradually building and supplementing the current systems with new data, the passenger component of ETIS can be developed over time.
The case study on passenger transport focused on data relating to transport between France and Italy. The intention was to test out how available data sources could be used to supply more relevant statistical information and to investigate the potential for harmonisation of quality, availability and suitability of combining national data of two countries. In the study it was discovered that data on French-Italian passenger flows were:

- of different quality,
- not easily comparable with one another, as they refer to data collections that represent different definitions, classifications, nomenclature and methodologies.

Despite computing efforts it was not possible to link figures to an O/D designation.

Transport networks (WP 6 and 7)

A desired transport network for ETIS should contain all the elements needed to represent transport infrastructure in a certain area such as roads, railways, rivers, air corridors etc. A transport network contains, beside the network elements, all the other elements needed for the representation of the real world in the system. In the case of MESUDEMO this relates in particular to the attributes of the transport network.

An extensive review of available European transport networks was carried out in MESUDEMO with respect to their attributes and which kind of geographical information system (GIS) that had been used. Pipelines were excluded because of the private ownership and the specific commodity flows.

The network attributes of various available GIS have been compared with the possible attributes from INFOSTAT and other European projects. In the INFOSTAT project 32 network node characteristics and 35 link characteristics were recommended. Concepts, attributes and characteristics can in the context of MESUDEMO be used interchangeably. In a very condensed form the options are summarised below. The recommendations from MESUDEMO are the following:

Road networks
The only way to guarantee an up-to-date accurate road network with many attributes is by using national sources. The national road networks are maintained in most countries on a continuous basis. Generally, it is recommended to:

- Define the ETIS-relevant roads.
- Define the geographical centres of NUTS-zones.
- Establish uniform definitions of attributes.
- Establish an organisation to receive the data from the national organisations on a continuous base.

A first step on the path to ETIS has already been taken by the RADEF project (see the following information box). This project contains a dictionary providing agreed terms, definitions, categorisations, and relationships for various types of road data, such as traffic details, number of accidents, road geometry and road condition statistics. Common location referencing is a particularly important component, linking the national networks to a European transport network. The data dictionary is the result of extensive co-operative research by the public road administrations. Further guidance from the European Commission to create an ETIS-network is, however, necessary.

Networks for other modes
Changes in the infrastructure for these transport networks occur less frequently. The recommendation is to continue the maintenance of existing network models on an ad hoc base. It is also recommended to establish an organisation to guarantee the quality of the network data. Furthermore, a standardisation of the definition of attributes is necessary. To extend the network concept to transport chains, the creation of a multimodal network is essential. Apart from the ETIS-relevant attributes also a unique transshipment location must be added to the separate networks.

It is recommended that compatibility between network and attribute data should be secured through using the network as the integrator of data. The only components that ETIS always must have are the maps of the transport infrastructure networks. There are substantial advantages of having only one common referential structure to which data can be related. Thus, some of the variables are location specific and need to be assigned as such. In this way many types of data sets can be related to each other through the spatial linkages and thus enabling complex relational analyses.

Available Geographical Information Systems (GIS) for transport infrastructure.

Infrastructure/ mode	**Networks described**	**Remarks**
Road	RADEF GISCO GISCO-APUR GEOSYSTRANS NIN United Nations Commercial networks (TeleAtlas, Mapinfo, NEA/IWW European networks and AND-mapping)	RADEF is a project from the Western European Road Directors in order to facilitate the exchange of highway information held in various national databases GISCO is the Geographical Information System of the European Commission GISCO-APUR is an update of the GISCO network by the French firm APUR. GEOSYSTRANS is a network made in the Fourth Framework Programme. NIN is the National Infrastructure Network of the Dutch Ministry of Transport, Public Works and Water Management
Rail	GISCO GISCO-APUR GEOSYSTRANS TERN	
Inland waterways	PAWN NEA PC-Navigo	PC-Navigo is a route planner for Western Europe's inland waterways developed and owned by a private Dutch firm. The NEA network is a network used in a production and assignment model on inland navigation predicting future traffic. PAWN is a network used for a model called 'Policy Analyses Waterway management for the Netherlands'
Sea	VESON	The VESON Worldwide Distance Tables is a privately owned database with static tables and routing along predefined linking points.
Air	GISCO	
Intermodal	GISCO Inventory of terminals Commercial: Ports database	

The Transalpine Pilot Study and its demonstrator (WP 10)

The Transalpine Pilot Study demonstrates the need and tests the feasibility of building up an ETIS to be applied to a specifically chosen area. The selected area should raise sensitive political issues at the same time as one might expect that the necessary data could be compiled and brought together. The choice fell on ETIS applied to the Alps. Therefore the database has been called the Alpine Transport Policy Information System (ATIS), and it proved to be a successful effort of combining trade and transport through the following steps:

- Regionalisation (NUTS 2) of trade flows between countries per transport mode and commodity group.
- Implementation of main transport chain structures in trade flows.
- Implementation of main Alpine transport chain structures in trade flows over the Alps.

- Distribution of trade flows along Alpine crossings over all main transport chain structures.
- Assessing impacts related to ATIS issues using the ETIS approach.

Before generalising the positive experience from ATIS to a future ETIS it has to be kept in mind that the data situation in the Alpine region is very favourable. Due to the focus of public interest in transport through the Alps many joint efforts have been made in monitoring these flows. Still there are other areas and corridors in Europe where comparable surveys, done in a harmonised way, would be very favourable for ETIS. Bearing in mind that in some of these areas national efforts of data collection have been made, the recommendation of harmonising and enlarging them on a European level looks quite promising.

A major problem faced in the Transalpine Pilot Study concerns data confidentiality. Many transport operators, mainly railway companies but also terminal operators are not willing to hand out their data in a sufficient degree of detail. Here solutions on European level are necessary.

A realistic picture of the important market of intermodal transport necessitates a transport chain approach within goods transport statistics. For that purpose a specific effort of data collecting should be undertaken. Although access to relevant parts of the information carried out by electronic data interchange systems (EDI systems) may be a solution in the long-term, the effort in the short-term has to be focused on shippers' surveys. As a complement, it would be valuable getting some specific information from terminals and ports.

ATIS used GISCO (the GIS used by the Commission) as its geographical information system, and included the ambitious aim of building a demonstrator to show how policy issues and the effects of policy instruments could be studied in the context of this geographic information system. Transport models reacting to selected policy instruments were taken from the STEMM project and adapted to ATIS. For this purpose so-called multiple route choice functions were calibrated for ATIS and data architecture and data models were developed.

As examples of policy instruments and challenges selected can be mentioned:

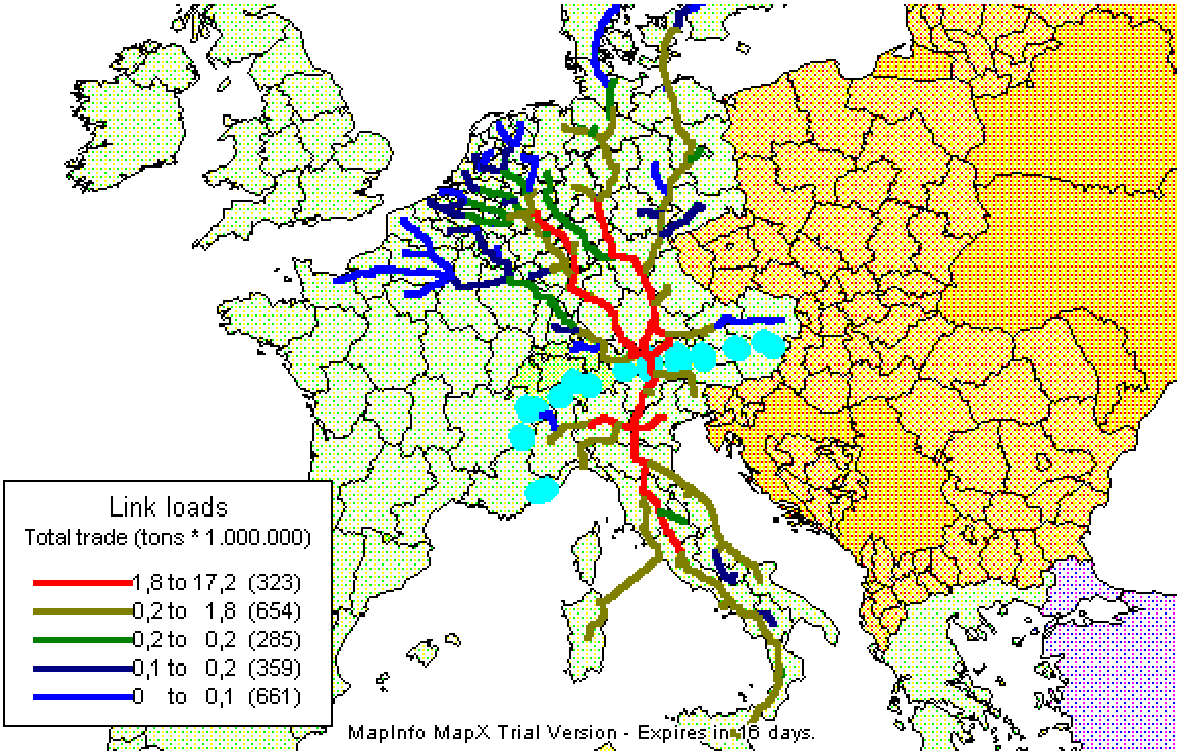
- Closing the Brenner Pass,
- Levying road tolls on vehicles crossing Switzerland.

The following two maps illustrate in the first instance road freight flows crossing the Brenner and the change to these flows if Brenner were to be closed.

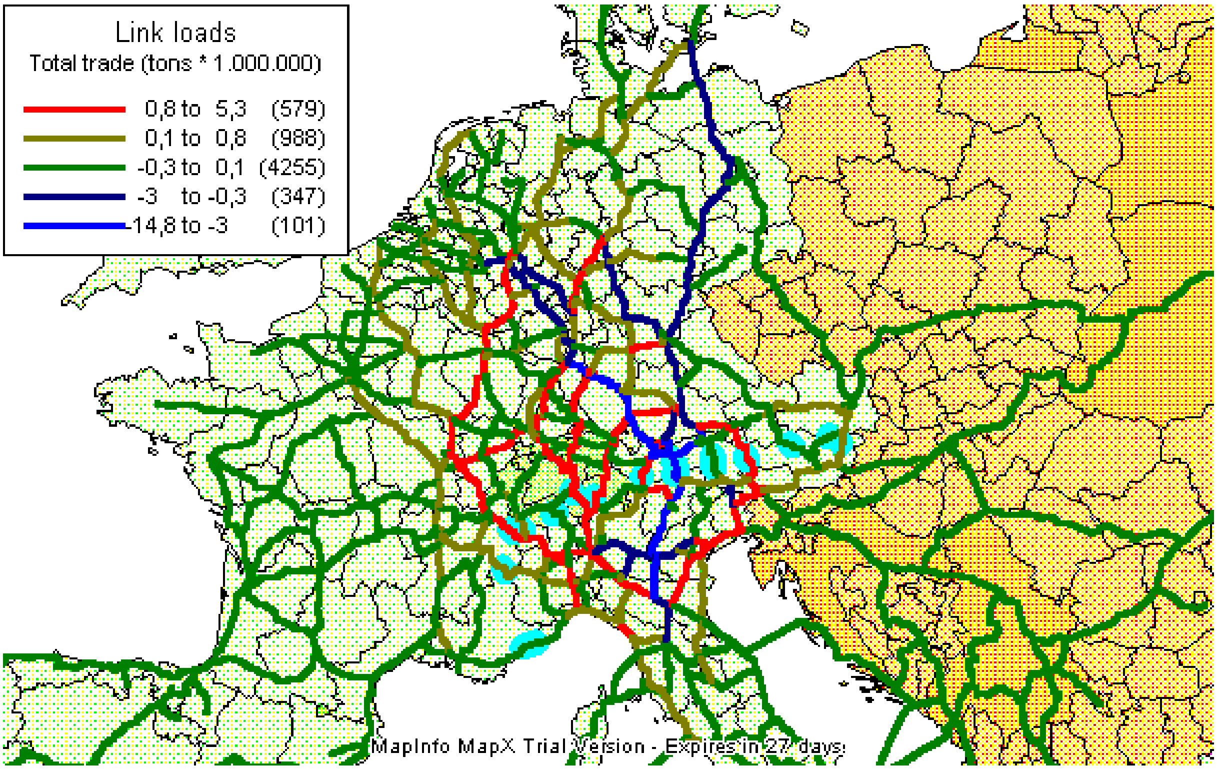
The implementation of ETIS has been revealed to be a complex project. The pilot study in the Alpine area was very useful in order to understand the process and to fix the steps of the implementation of a future ETIS. Particularly the possibility to build up a European database for freight flows has been highlighted and is an important added value of the work.

The final lesson is that the architecture and the structure recommended for ETIS have been proved to be feasible. It is also recommended that the architecture should be built *by default* to begin with, as most of the options are fixed, and to extend it in order to introduce more flexibility in the measurement and analysis of policy impacts.

Structure of road freight flows crossing the Brenner



Change on road freight traffic flows when closing Brenner



Ensuring compatibility (WP 9)

Compatibility between supply and demand methodologies is one of the universal goals for a system that will monitor and guide policy actions in transport. Such compatibility will not only make it possible to provide the user with information on the actual utilisation of the transport infrastructure, but also increase the knowledge of the forces behind it.

When the objective of MESUDEMO was formulated, it was presumed that methodologies for supply and demand databases would be developed during the MESUDEMO project work, leaving the finalising stages "only" with the challenge of ensuring compatibility between the two. However, during the work many other compatibility obstacles surfaced, as difficult and as important. The compatibility challenge, then, should be viewed along two dimensions, one between already existing data and systems, and the other between how mobility is conceptualised and data selected and compiled to describe and explain it. The figure illustrates the two dimensions.

The two-dimensional compatibility challenge

MESUDEMO mainly deals with the challenge of ensuring compatibility between existing systems, but the other approach should not be forgotten: A dialogue on the forces behind transport should be encouraged in the European countries, leading to more compatible systems underlying ETIS.

Data architecture and data models for ETIS (WP 4 and 8)

The main objective has been the development of a set of data models and the corresponding data architecture for the creation of a European database on transport flows and transport infrastructure.

The main results of WP 4 and WP 8 can be summarised as follows:

- A better understanding of the organisational and technical nature of ETIS has been achieved as to aim and form, technical needs and solutions, installation and maintenance etc.
- A step-wise practical path to a future ETIS has been envisaged and described.
- A concrete need for preparatory work at the European level has been identified.
- A general architecture of an ETIS-type system has been developed, and the main modules of it have been further analysed.
- A novel internal data architecture and the corresponding data model for the transport infrastructure networks and the transport flows have been developed, which:

  - For the first time, takes into account critical parameters such as the level-of-detail, the relevance of the associated information (e.g. the network and region attributes) to the transport or policy issues at hand, the possible locality of the problem at hand etc. (e.g. the local problems of specific corridors)
  - Can support widely different needs, ranging from organising transport supply and demand information to transport modelling and forecasting. It can also resolve problems from the polymorphy and diversity of the transport data.
  - Permits the seamless integration of external, third-party data sets and processing methods (including models), without violating ownership and confidentiality issues.
  - Follows the current state of art for creating open and scalable systems.
- An implementation methodology for ETIS that reduces the installation and maintenance costs and permits the open, but controlled access to its data and processing methods, has been developed. This methodology, moreover, permits the easy construction of third-party additional software for ETIS.

The following figure outlines the general architecture of ETIS, showing the various sub-systems at a conceptual level.

ETIS will comprise data of two categories; structural and unstructural data. The first category can be further divided:

- *Static data*, i.e. data that regularly do not change or change very slowly. Usually most of the infrastructure-related data as well as socio-economic data fall in this category.
- *Dynamic data*, i.e. data that change very fast and need special methods for handling, e.g., traffic data on a road network that may need updating very often.
- *Time series* comprising the time pattern of indicators.

Unstructured data are usually textual or multimedia data, including documentation, EU or national directives, knowledge representation, etc.

Usually, a Geographical Information System (GIS) is used for the organisation of the infrastructure-related data. The transport networks, specifically the nodes and the links, are used as the elements, to which all other information is linked. Each mode is usually on a different plane, called layer. The work in MESUDEMO, however, has shown that such a conventional organisation can prove both inefficient and restricting for ETIS.

In MESUDEMO, a unified model has been developed for ETIS, for handling all different kinds of data, as well as the necessary meta-information. As a general rule, all data are recommended to be organised in a contemporary, advanced DBMS (Database Management System). The need for geographical visualisations and spatial analyses are covered through the use of a spatial data interchange module, which takes care of the communication of the database to a GIS. Such modules have already become commercially available.

A general architecture of ETIS, showing the various sub-systems at a conceptual level

ETIS must be associated not only with an information system, but also with a whole set of methodologies and procedures for gathering proper data, creating and regularly maintaining a corresponding information system suitable for supporting decision making in transport policy matters.

In the case of ETIS, the usual top-down design method, used for national systems cannot stand alone. Instead, a combination of this approach with a bottom-up one has to be used. It may be necessary to build specific, localised ETIS sub-systems for dealing with problems and political issues of multi-national character particular for certain areas. These systems can be thought as a more complete and extended form of transport observatories. A possible and probable form that an ETIS can take is that of a network of interconnected, co-operating (i.e. federated) systems: **the ETIS network**.

One must carefully consider not only the initial creation of ETIS, but also the maintenance and update of it as major challenge. In fact, building a complete information system of this type is only justified, if this system is highly re-usable and updateable.

INTRODUCTION

The Common Transport Policy (CTP) is given high priority in the European Union (EU). The rapidly increasing mobility observed for many years has consequences like bottlenecks in the infrastructure, damaging impacts on the environment and regional imbalances. There is a need for strategic transport research in order to provide relevant knowledge to support short-term corrective policy actions as well as to support long-term strategic policy actions aiming at structural changes.

Main problems in the transport sector are:

- economic inefficiencies, including increasing transport costs and the inadequacy of the transport networks to cope with the growing logistic requirements and mobility needs in relation to commercial competitiveness
- increasing congestion problems in the central parts of Europe and in larger metropolitan areas
- an intolerable large number of traffic accidents and related high costs to society in general
- increasing environmental damage and growing energy consumption
- geographical imbalances and inefficiencies accented by the present transport network structure.

Road transport is dominating and congestion and externality problems create increasing pressure on policy makers to induce shift to rail and shipping and various forms of intermodal transport. The railways have difficulties to respond to increased expectations, not being competitive, neither from the cost side nor in terms of providing competitive quality of service in a modern logistics context relevant for shippers. Infrastructure links across Europe are located in different countries with different policies and regulatory traditions. The main problem regarding rail is that the presence of national monopolies with an inadequate internal organisation and a lack of market oriented behaviour hinder competitive strategies. This is reinforced by the fact that railways have to serve national policy goals where transport development *per se* is sub-ordinate to regional development, etc. Similar "European" challenges have been pointed at in air transport. In this sector it is claimed that there is a potential for a more efficient hub-and-spoke organisation in Europe than the present one, which has been motivated to a considerable extent by the national orientation of airline companies, with national capitals as hubs, regardless of other alternatives.

Looking at the policy itself, initiatives are characterised by the dominance of infrastructure concerns. But "hardware" in general is not a scarce factor, and hence constructing new links, tunnels, etc., cannot be the crucial success factor. Large infrastructure improvements will have impact in the long run, and in a way only create potentials, but they will not by themselves improve the efficiency of the system.

The problem of promoting intermodal transport contains many features that are typical for European transport policy. Being a problem of intermodality, the success of combining transport modes also requires solutions to problems of interconnectivity and interoperability. The relevant transport distances being long, more than one country is involved. Therefore infrastructure planning (including terminal facilities), regulations of the use of these infrastructures (access, tariffs, quantitative regulations) and subsidies to promote intermodal transport have to be co-ordinated among different countries with different interests. Promoting intermodal transport in an increasingly competitive and liberalised market also implies considering behaviour on the supply side, but above all the demand side. The weight also has to be on promoting efficient use of infrastructure rather than building infrastructure.

An important component for the CTP is to create conditions for the interoperability and intermodality of transport networks so as to achieve the most efficient and cost-effective use of the different modes of transport, having regard to their effects on safety, environment and energy requirements. In the *Transport White Paper* the objectives and scope of the CTP are described, in relation to the principle of sustainable mobility:

> *These objectives require the development of policies to ensure that the transport sector can take full advantage of the Treaty's provision on the single market and that the different geographical components of the Community benefit from transport systems that will provide services efficiently, safely, under the best possible social conditions and fully respecting the objectives of the Community's environment policy* (paragraphs 35 - 39).

Trans-European Transport Networks (TEN-T, or here shortened to TEN) aim at meeting the mobility requirements of both citizens and enterprises at reasonable cost consistent with their long-term maintenance and development. The policy objectives regarding environment indicate that transport systems must contribute to the solution of major threats for the environment, such as the greenhouse effect and to the achievement of sustainable development.

The objectives of transport efficiency and social cohesion have a spatial dimension as well, in terms of meeting regional development goals and considering the spatial distribution of socio-economic activities, in particular in relation to the development of the TEN. Globalisation together with an open and liberalised single market creates challenges of a new kind. *European Spatial Development Perspective (ESDP)* is a political tool to reduce regional differences in growth rate and living conditions. The CTP objectives relate to the problems derived from the increasing mobility and the expanded needs for mobility at all geographical levels and in all sectors of the society.

Community research concerning transport should provide new tools for facilitating sustainable mobility: efficient, safe transport under the best possible environmental and social conditions. Internalisation of external costs by pricing or taxation of each mode and system is envisaged to be a major element of the CTP.

Regarding the principle of subsidiarity it is stated that the EU and its Member States have to co-ordinate their research to ensure that national policies and Community policies are mutually consistent.

These, and similar, policy problems at the European level call for tools to measure and monitor transport performances in a proper way, to link the transport sector to the other sectors of the society, to be able to select the most efficient actions and to calculate - and predict - the consequences.

The important process of developing and diffusing a unified European transport policy that appears credible to all European countries, acknowledging that the shape and content of mobility patterns have evolved out of different social and economic histories, also has its imperatives on transport research. While the European society becomes more economically, socially, and politically integrated, this to an increasing degree underpins the geographical dispersion of activities. Social and economic processes are functionally more integrated than before, but at the same time played out over an extended geographical arena: Local problems do not necessarily have local causes and general solutions to specific problems may create adverse effects in different localities.

To a project like MESUDEMO, the aim has been to develop sets of methods that will cater for decision support for European transport policy design. This challenge has boiled down to two main concerns: Firstly, there is a need to harmonise the generation of transport information across Europe in order to achieve a consistent description of transport development. Secondly, the generation of such information must be designed in a way that promotes an enhanced understanding of the forces that shape and drive transport development. These are both formidable tasks, and the quantitative and descriptive orientation that was a precondition for MESUDEMO has lead to a focus on the first point.

The challenge of promoting knowledge about the efficiency of infrastructure utilisation resulted in an organisation of MESUDEMO along two lines. One line treats data on infrastructure provision, the supply side, and the other deals with data on transport flows, which then constitutes the demand for infrastructure in this context. Extracting, processing, managing, and disseminating such information at the European level through the use of state-of-the-art information technology, is a major achievement of the project.

According to the main lines sketched above, MESUDEMO was originally designed as parallel strings of work packages, with a compatibility check all along the project work and finally in the end. As often experienced, however, knowledge does not generate in such a one-dimensional, linear fashion. Just as much, experiences from later work packages influenced on preceding ones, spurring further work there. This in turn laid the foundations for further work in the following work packages, and so forth. This experience has necessitated a dynamic way of working.

The non-linear, or iterative form of knowledge generation has presented the project with a series of hen-and-egg problems and solutions have not always been inevitable: How does one design transport information for the support of policy decisions when these decisions are based on transport information? How does one, technically, design a transport information system when the content of these databases depends on what is technically viable?

In addition to this, the project has been exposed to several “top-down versus bottom-up” considerations in different situations: A re-design of European transport information fitted exclusively to the quest for decision support on the European transport policy level would face the problem of possibly disrupting both the structures and content of existing transport information systems. This would be a costly process, while leaving much existing knowledge obsolete. On the other hand, a future ETIS could be built on highly valuable, though fragmented and less compatible, existing systems and data, but then possibly limiting the scope for decision makers on the European level in designing adequate policies.

From the outset of the project, it was a common opinion within the Consortium, as well with representatives of the Commission that the recommendations arrived at by “desk research” would not be sufficiently grounded on the abstract level alone. This led to an extension of MESUDEMO in shape of the Transalpine pilot study. This extended the scope of the project to a considerable degree, and was an opportunity to combine knowledge from several other 4th Framework Programme research projects. Furthermore, it was invaluable for testing the geographically generalised conclusions of the project in a specific context. This is an important concern as the intertwined relationships that shape transport development make such development highly context specific. It was therefore regarded as important to test the relevance of the generalised MESUDEMO priorities in a specific area. This generated invaluable knowledge both for the relevance of the MESUDEMO recommendations and for further work. It should be remembered, however, that such a case study is considered more like an experiment, less like a “sample” prone to generalisation. The potential of case studies for generalisation lies in the possibility of expanding theories (analytical generalisation), not in enumerating frequencies (statistical generalisation).

Transport research must respond to policy priorities, but as the political perception of transport also is coloured by deliveries from transport research, this is not just a one-way relationship. A gradual evolving and refining of the perceptions of how the transport system develops hinge on interrelationships between decision makers, planners, researchers, and other actors within the system. Therefore, it has been a main concern in the project to establish MESUDEMO as an arena where such mutuality could be reached.

MESUDEMO has been a large project, in resources deployed, in time-span, and in thematic scope. Still, it cuts into real life only at a certain stage of development, which underpins the need for continuity in transport research compatible with the continuing and ever-changing character of society itself.

It has been argued that all priorities may not be simultaneously achievable, nor do they exist independently of each other. There may also be discrepancies between some kind of scientific ideal and practical research, where beliefs, doubts, trying and failing are the everyday experience. However, these issues do find their solutions through human practice. How these priorities were weighted, manifested itself through the actual selection of methods in the project. The Consortium, mainly with experience in research and consultant activities, was strengthened in this respect by the active commitment and involvement from the national administrations participating in the project.

This report presents the methods, the work process and the findings in MESUDEMO in more detail; a more comprehensive description of the information can be found in the Deliverables of the project, listed in the Partnership chapter above.

Chapter 1 explains the methods applied in MESUDEMO, from the Questionnaire survey through the in-depth interviews, to the pilot studies in the Alps and in the Netherlands and the two Workshops arranged by the Consortium.

In chapter 2 the recommendations on how to represent transport flows on freight and passengers are summarised. In chapter 3 we present our recommendations towards the networks to include in ETIS, in order to design a simple and flexible decision support tool that integrates transport flows and the infrastructure supply. Chapter 4 covers the Transalpine pilot study that was performed in the framework of the MESUDEMO project, all the way from policy issues through data acquisition and combination, to the construction of a harmonised database for freight transport chains crossing the Alps.

Chapter 5 deals with the challenge of ensuring compatibility between supply and demand data. Two different approaches are introduced. One way of ensuring compatibility is by doing the best out of existing data, by harmonising and organising data from different sources into consistent data sets. The other approach is to encourage a dialogue between the suppliers of data on how we conceptualise transport development, thereby securing compatible systems from the bottom of ETIS. Chapter 6 describes the architecture of ETIS and the data models that will secure a flexible and consistent system for European transport policy decision support. The final chapter 7 presents the overall conclusions and recommendations from the project.

1 RESEARCH METHODS SELECTED FOR MESUDEMO

1.1 Framework

The project was organised according to ten different workpackages (WP's), interlinking each other as shown in the figure below.

Figure 1-1:The Workpackages and organisation of MESUDEMO

The method chosen for the MESUDEMO project can be described in a three-stage approach:

1. Make an analysis of significant discrepancies between existing procedures used to produce current databases and their associated data variables. Study the transition from the databases, which hold national information to those, which have to hold European, multi-national information.
2. Describe and develop viable, practical methods fitting the basic concepts of the INFOSTAT project, and draw up the presentation framework for the creation of European databases.
3. Test the method for constituting and presenting the database by using existing and/or new, measured or modelled data.

MESUDEMO started off with WPs 1 and 5, stage 1 above, with a questionnaire survey and in-depth interviews mapping the state-of-art in Europe when it comes to transport databases and information systems. In WPs 2 and 3, stage 2, suggestions were developed towards variables on flows that were essential as the basis for current and future policy decisions. These workpackages also supplied a procedure for estimation of those variables that were urgently needed, but missing at the time. The network to include in the European Transport Policy Information System (ETIS) was developed in WPs 6 and 7, stage 2, both for modelling and for policy decisions regarding infrastructure. In WPs 4 and 8 it has been proposed an integrated architecture for the system that revealed itself through WPs 1-3 and 5-7; an overall system architecture, internal data architectures for the network and flows data, as well as proper implementation methods. Finally, in WP9 the compatibility between the supply and demand data and tools has been considered. The Transalpine pilot, WP10, along with the pilot in WP 2, were both laboratories for the findings developed in MESUDEMO, but contributed also as a first stage on the path towards implementing ETIS (stage 3).

1.2 Research orientation

In the MESUDEMO project a need was felt for a further step into reality, trying to "get behind" the information gathered in INFOSTAT. The original main purpose of a questionnaire survey was twofold:

1. Since MESUDEMO was going to come up with methodologies for a complete information system, it was regarded essential to have good descriptions of alternative ways of organising transport information systems around Europe. We thus wanted to identify some good practice to learn from, in the process of describing the architecture of

the pan-European system. This called for a comprehensive survey of each database, getting insights into every step and functionality in the production process of the database. At the outset we did not have sufficient knowledge of the distribution of such good practice in Europe and thus we wanted to start out broadly, trying to get a large number of answers to a large number of questions. This initial approach had the clear disadvantage that we risked having a low response rate, since filling out the questionnaire could be a time consuming burden on several persons in each institution.

2. The second purpose of the questionnaire survey was found in the fact that many of the respondents represented possible partners when building the architecture of ETIS, in the form of sources of information. Therefore it was regarded as interesting to compare and assess the possibilities of linking the respondents together in a future information network.

The sample was decided to be as wide as possible in order to create a first linkage among MESUDEMO and key actors fully involved in MESUDEMO outcomes. Thus, the need to reach as many as possible database holders resulted in a sample of 99 respondents across Europe. Some relevant respondents were found among the institutions that received the INFOSTAT report as part of the dissemination activity of the project. Also, the list of database holders from the UN/ECMT/ EUROSTAT common Questionnaire was used as a basis for revealing good databases.

The questionnaire contributed to the project in several ways:

1. It was a useful process to raise awareness of the whole MESUDEMO project, thus the construction of the questionnaire represented a starting point for the work on all work packages.
2. The results from the survey constituted a basis for WPs 1 and 5, describing existing methods for transport information systems in Europe.
3. The general impression from the responses was a necessary departure point for selecting interview objects for in-depth interviews later on in the process and for establishing contacts with transport database holders.
4. The survey provided MESUDEMO with inputs about possible solutions for a Pan-European database on infrastructures and of flows. The questionnaire was to give inputs both to how an information system for transport could be designed, that ETIS could learn from small systems, and also on how these small systems could be integrated into ETIS.
5. The broad distribution of the survey in itself established attention towards ETIS among experts in the European countries.

The drawback of a quantitative questionnaire survey as a method, is the lack of contextual information. In order for MESUDEMO to provide a vision of ETIS, we needed experience on the complete production process of the database. The respondents' different preferences on the ”modules” in the system (surveys, harmonisation, software, hardware, dissemination media etc.) are probably better explained by the context they belong to, in time and space, than the isolated characteristics of the selected solutions. It would thus be a big mistake if we did not go deep into selected cases, but just selected the most frequently used solutions to the ”modules” needed in ETIS.

The advantages of investigation through in-depth interviews are that one is more able to get a grasp on the functionality and organisation of a complete system. In addition, in the course of the interview, one can detect aspects not contemplated before the investigation. A drawback is that interviews of this type are rather resource demanding. Thus, only a limited number of database holders were interviewed.

The aim of the interviews has been twofold. Firstly, to get a better idea of the state of the art when it comes to collecting, storing, maintaining and disseminating transport information. This has been an important part of the input necessary for the other work packages that seek both to come up with a methodology for the creation of ETIS and to make a small-scale exercise of what ETIS might look like in practice.

Secondly, the goal has been to establish a network of data providers in Europe and to promote the idea of ETIS to them. The success of ETIS relies heavily upon co-operation with data providers and their attitude towards the idea of establishing and maintaining ETIS.

The questionnaire survey and other inputs were the main basis for the selection of institutions to be interviewed. Some international database holders quickly pointed themselves out, but other sources were contacted as well. The selected institutions included national statistical offices, transport ministries and other public institutions, supranational institutions and commercial companies.

In total 12 interviews were carried out in 5 countries. The organisation interviewed comprised three international organisations (EUROSTAT, ECMT and UIC), two national statistical offices (Italy and Norway), three Ministries of Transport (the Netherlands, Italy and France), one port (Rotterdam), one railway company (Italy) and two motorway companies/ organisations (Italy and France).

Workshops were considered essential for co-ordinating project work as well as generating ideas. The workshops were both instruments for scientifically progress *within* each on-going work package, as well as a first introduction of the coming WPs. The workshops also functioned as an instrument for disseminating acquired knowledge between each WP, as well as between the MESUDEMO project and related projects and institutions.

The first workshop was arranged in Copenhagen on May 11th and 12th 1998. In addition to the members of the Consortium participated EUROSTAT, UN/ECE, Dutch Central Statistical Bureau, Ministry of Transport in France and Ministry of Transport in Great Britain (DETR). EUROSTAT participated after that in many MESUDEMO meetings. The workshop produced many valuable insights into the common questionnaire, and into transport databases in Europe.

The second Workshop took place in Rotterdam on the 17th and 18th of June 1999. The presence of several representatives of the Commission, producing inputs towards the needs for ETIS and the role of MESUDEMO in this context, was one of the main benefits from the second workshop. External experts, both from other 4th Framework projects (ASSEMBLING, CONCERTO, BRIDGES), but also other experts gave valuable inputs to the work in MESUDEMO.

For the sake of dissemination and contact with relevant institutions and experts, we created our own homepage, where the project and its results are presented.

1.3 Pilot studies and test cases

It is a widely accepted technique to execute pilot studies in order to get a better insight inotcomplex systems which are difficult to grasp from the beginning. The pilot studies and tests in MESUDEMO served at least three purposes, or one purpose on three levels:

Firstly, "semi-experiments" were needed to test out hypotheses developed in the project. The real-world examples were hard to access. Also, there were few examples of organising information geared towards pan-European policy issues. Existing systems are often either national or uni-modal.

Secondly, the questionnaires and the in-depth interviews did not give sufficient unambiguous answers on the feasibility of the vision of the architecture of ETIS. It was necessary to check results from desk research and meetings against limited problems and geographical areas in Europe. MESUDEMO was going to recommend methodologies for establishing ETIS, and the pilots were used to check to which degree results were feasible in the European transport environment.

Thirdly, the pilot studies would produce knowledge about what could be done with *existing* data and models, and, thus, if it was worthwhile spending resources on making a pan-European system based on existing data and tools.
It would not be possible to generalise conclusions to the whole of Europe based on the experiences from each pilot. However, the pilot studies have led the Consortium closer to general knowledge.

1.3.1 The Transalpine Pilot study

The idea of the Transalpine Pilot study was launched at the Paris workshop of CONCERTO in May 1998. The transit transport in the Alps was considered a good example of a policy issue with a strong European interest, and a proper one to be used in a pilot study for testing the feasibility of a system consisting of elements of an ETIS.

The Transalpine Pilot was aiming at examining in practice and in detail the way to use transport knowledge and existing data sources and models for answering policy questions. The aim was to construct a system with many of the same characteristics as ETIS.

The Transalpine Pilot was connected to a complex transport area with a high degree of political interest, where four EU countries and one non-EU country (Germany, France, Italy, Austria, Switzerland) are directly involved, in addition to a host of other more remotely involved countries. There is a high degree of Community interest negotiating and regulating Alpine transit, also trying to promote sustainable environmental friendly transport solutions. A huge volume of freight crosses the Alps, raising questions about mode choice, the use of existing infrastructure, environmental impacts etc. Several studies had been undertaken corresponding to different interests, policies and data problems.

The main objectives of the Transalpine Pilot were to investigate how far towards ETIS one could come by using all available data and information and to demonstrate the *need* of an information system.
The Transalpine pilot was defined as a separate WP10 in MESUDEMO, thus replacing the need for a separate pilot study in WP9.

1.3.2 The TRANSITIE Pilot study

In the work of WP2, aiming at the recommendation for variables on freight flows, the TRANSITIE pilot was performed in order to check whether it is possible to get a consistent and reliable image of the international goods flows from, to and through a limited case area. In the Netherlands needs were felt for reliable data on international goods flows. If no action would be undertaken, both quality and quantity of the statistics would deteriorate. Therefore it was desired to check the feasibility of getting this consistent and reliable image of the international goods flows from, to and through the Netherlands by other means. Statistics of import and transport was an important source of information for the international part. In the pilot an attempt was made to get better insight in the bottlenecks which were observed in the various statistics for international goods transport and the possible solutions.

1.3.3 Mirror bias error: An example on France/Italy passenger flows

The example from WP3 focused on the passenger transport flow between Italy and France. The method was firstly to identify and contact data holders for each mode in each country. The data holders are both public and private bodies, situated at national, provincial and local levels, and indeed covering different territories. Therefore, a second task was to collect the relevant data and control if the data fitted the required specification or could be modified to fit into the definitions of the selected variables. The third step in the method was to check the national data in the two countries on quality, availability and suitability, and in addition on the technical computer side and on the process to acquire the data.

1.4 Methodology for the work on the ETIS Architecture

The design process of any information system must start with the needs of the users of the system. From their needs one can deduce the functionality, the hardware and software needs and the quality of the data to be fed into the system. In an information system like ETIS this has not been possible. There are many reasons for this, one of them being that instead of the needs of the system guiding the data provision process, we depend on existing data gathered in the European countries. Another reason is the uncertainty regarding the needs of the users of ETIS. Being a social system, it is extremely challenging to organise the constituent parts of the transport system into an information system that will reflect the causes and effects in transport development. This objective is very ambitious, and may be impossible to meet completely.

The following methodology was initially planned for the ETIS architecture:

- State-of-art analysis
- In-depth analysis of existing “best-practice” databases
- Consideration of the sources around Europe, in order to find a way to create ETIS as a virtual database (a concept first presented in INFOSTAT)
- Determination of the requirements and, if possible, the specifications for such a system
- Proposition of an overall architecture and internal data models for the flows and infrastructure databases at a European level
- Getting feedback from the potential users of the system on the proposed schemes
- Execute selected computer tests to demonstrate the feasibility of the proposed schemes

- The work plan was necessarily updated each time new State-of-art information was available and/or evolving databases became available for examination. Moreover, based on a continuous monitoring of the State-of-art and the evolution of the policy and transport issues and problems, an extensive analysis of the concepts relative to ETIS has been made and an additional, important set of design principles has been identified.

The methodology has been complemented with the following steps:

- Broadening the State-of-art analysis, in order to include the international State-of-art (not only in EU, but also in North-America, etc.) in advanced transport information systems and systems including georeferencing
- Examination of the possibility of technology migration from other IT fields
- Investigation of the possibility of creating open and scalable architectures, as well as systems that can:
  - Accept and handle the diversity and polymorphy of the data and processing methods used for the transport-related and policy-related problems
  - Cope with issues like data ownership, locality of the problems and the policy issues, etc.
  - Achieve a graceful co-operation with the multitude of existing national and international systems.

2 TRANSPORT FLOW INDICATORS

The term *indicator* is widely used in the INFOSTAT project. However, INFOSTAT that introduced ETIS used the terms *indicators* and *variables* rather interchangeably. One might argue that indicator is a wider term which is given a more quantifiable content by one or more variables. The quantified content is expressed through data and statistics. One might also argue that variable is a wider term than indicator. The ranking does not matter in MESUDEMO and, therefore, no strict ordering has been observed. Mostly the term indicator is used in MESUDEMO.

Transport takes place in space and thereby creates time and place utility. The actual division of space into regions or zones, two concepts that are used interchangeably in MESUDEMO, might be crucial to the kind of detailed information, which ETIS can supply. Selection of an actual zoning system to be applied to ETIS has not been specified as a task for MESUDEMO.

An administrative classification system of space called the NUTS system (Nomenclature des Unités Terrioriales Statistiques) has been developed by the EU Commission and has been used in MESUDEMO to illustrate the spatial reasoning.

The NUTS system is specified for several spatial magnitudes. In NUTS 0 the zone is equal to the territory of the national state. So an O/D specified for NUTS 0 means that the whole of the country is considered one area. Zones with higher numbers represent smaller areas within each country. For each country the NUTS level and zone boundaries have been defined and are known.

The basis for delimitation of zones has - in most cases - been the existing administrative borders. Hence, the size of zones in one country may vary considerably compared to other countries. Another point to note is that the zone borders in many cases will not be the best delimitation for the purpose of accumulating transport information. To use the case of transport modelling, the zone system should ideally be based upon functional criteria and not on administrative borders. In practice a compromise between planning needs and availability of information has to be reached.

When ETIS has been fully developed at some NUTS level, it will of course be possible to aggregate up NUTS levels according to the problem to be solved or the question to be answered. Specification for more details might be impossible when the NUTS level of ETIS once has been set.

NUTS 3 is a zone size at which many of the policy questions asked can be answered at a sufficient level of accuracy and detail and NUTS 3 is also the level at which several nations have chosen to develop their national transport planning models. In practice NUTS 2 might be the best to aspire for given the present state of affairs in databases.

2.1 Goods transport indicators

2.1.1 ETIS indicators and the need for a consistent database

The ETIS indicators as developed in the INFOSTAT project have been at the core of WP 2. To test out the proposal for operational indicators and in order to discover the difficulties that exist in practice to arrive at ETIS, a pilot study was carried out in WP 2. This was achieved through co-financing with the Dutch Ministry of Transport and with data on international freight flows from the Netherlands. This pilot was called TRANSITIE.

The work developed in four stages:

1. Developing a definition of relevant indicators to constitute the database on goods flows.
2. Development of a methodology for a new statistical framework with data from the Netherlands: The pilot study TRANSITIE.
3. A pan-European extension of the methodology for a general database on goods flows.
4. Recommendations and inputs to other WP's of MESUDEMO.

The specific fields of investigation are the goods transport demand indicators of ETIS that are needed to describe the volume and structure of transport generated by the different mobility actors. The selected indicators are listed below in a uniform way, including the label, the unit of measurement, observational unit and priority level in accordance with the INFOSTAT recommendations.

None of the indicators may be considered as being final, neither in INFOSTAT nor in MESUDEMO. They may be elaborated further in future, but they are now brought into a harmonised and consistent database structure, using the experience gained in the TRANSITIE pilot study on the data from the Netherlands.

The operational definitions leading to a policy oriented comprehensive database structure are described in table 2-1.

Table 2-1: ***Refined INFOSTAT goods transport demand indicators***

No.	Label	Unit of measurement	Observa-tional unit	Priority level[3])
1.	Total annual interzonal goods transport flow by commodity group, mode (or combination of modes) and type of transport chain	Tonnes/year[1)]	O/D pair	F
2.	Average distance between origin and destination of transport unit by mode	Km,	O/D pair	F
3.	Average distance between origin and destination of the good	Km	O/D pair	F
4.	Loading factor (ratio between volume and capacity) per type of transport unit by mode	Per cent	O/D pair (main leg)	D
5.	Annual total number of tonnes transported broken down by • size of shipment (weight) • value of shipment[2)] (ECU) • trip distance (km) • containerisation (yes or no) • type of transport unit	Tonnes/year[1)]	O/D pair (main leg)	F
6.	Average number of days of use of the transport unit	Days	Zone	D
7.	Use of EDI, tracing and tracking of shipment	Yes or no	Zone	D

1. Tonne-kilometre figures may be derived when combining the tonnes/year and distance between O/D pairs.
2. It might be considered to present the value flows in the same spatial and modal details as volume (tonne) flows.
3. F = fundamental, D = desirable

Indicator 1

No.	Label	Unit of measurement	Observa-tional unit	Priority level
1.	Total annual inter zonal goods transport flow by commodity group, mode (or combination of modes) and type of transport chain	Tonnes/year	O/D pair	F

This indicator is intended to describe the volume and structure of goods flows between the original origin and final destination. Following the INFOSTAT definition transport chain has been defined as: "*a sequence of transport modes used to carry a certain quantity of goods from its origin to its final destination. Along the chain, one or more transhipments may take place. As such direct flows of goods may be regarded as a transport chain without transhipment.*" (Page 46)

A transport chain with a single transshipment node (point in space) could be represented in a database record structure as follows:

1. *Zone of original origin of goods flow (o)*
2. *Zone of final destination of goods flow (d)*
3. *Zone of transshipment (n)*
4. *Mode from original origin to node of transshipment (m1)*
5. *Mode from node of transshipment to final destination (m2)*
6. *Commodity group (g)*
7. *Weight transported in tonnes (TT)*

In mathematical notation: $TT_{o, d, n, m1, m2, g}$

Indicator 2

No.	Label	Unit of measurement	Observa-tional unit	Priority level
2.	Average distance between origin and destination of transport unit by mode	Km	O/D pair	F

This indicator could be interpreted as the infrastructure network distance with choice of shortest, cheapest or fastest route between two regions. Transport time and cost could also be calculated in a similar way. The term transport unit is explained in table 2-2. Indicator no. 2 could be represented as follows in a database record structure:

1. *Zone of loading of transport unit (i)*
2. *Zone of unloading of transport unit (j)*
3. *Transport mode (m)*
4. *Transport unit (u)*
5. *Average distance in kilometre (DT)*

In mathematical notation**:** $DT_{i,j,m,u}$

If distance is calculated from available O/D-data in transport statistics on both weight transported (tonnes) and transport performance (tonne-kilometres), transport distance could be represented in a database record according to:

1. *Zone of loading of transport unit (i)*
2. *Zone of unloading of transport unit (j)*
3. *Transport mode (m)*
4. *Transport unit (u)*
5. *Weight transported in tonnes (TT)*
6. *Transport performance in tonne-kilometres (TK)*
7. *Average distance in kilometres (DT) = 6/5*

In mathematical notation**:** $DT_{i,j,m,u}$

$$DT_{i,j,m,u} = \frac{TK_{i,j,m,u}}{TT_{i,j,m,u}}$$

Indicator 3

No.	Label	Unit of measurement	Observational unit	Priority level
3.	Average distance between origin and destination of the good	Km	O/D pair	F

Just as for indicator 2, it seems justified to relate this indicator to transport performance in a transport chain, which could be calculated from available O/D- (chain-) data on tonnes transported and tonne-kilometres. For a transport chain with a single transshipment node the average distance could be represented as follows in a database record structure:

1. *Zone of original origin of goods flow (o)*
2. *Zone of final destination of goods flow (d)*
3. *Zone of transshipment (n)*
4. *Mode from original origin to node of transshipment (m1)*

5. *Mode from node of transshipment to final destination (m2)*
6. *Weight transported in tonnes (TT)*
7. *Transport performance in tonne-kilometres (TK)*
8. *Average distance in kilometres (DT) = 7/6*

In mathematical notation**:** $DT_{o,\ d,\ n,\ m1,\ m2}$

$$DT_{o,d,n,m1,m2} = \frac{TK_{o,d,n,m1,m2}}{TT_{o,d,n,m1,m2}}$$

Indicator 4

No.	Label	Unit of measurement	Observa-tional unit	Priority level
4.	Loading factor (ratio between volume and capacity) per type of transport unit by mode	Per cent	O/D pair (main leg)	D

Although this indicator is not seen as fundamental for the transport demand flows, it could give valuable information on the efficiency of transport performance both for transport operators and policy makers. The capacity of a transport unit can be specified in various ways, for instance:

- *Weight carrying capacity (tonnes)*
- *Volume carrying capacity (m^3 or litres)*
- *Capacity in number of units (TEU's, pallets, length of lanes, etc.)*

A database record with the loading factor (capacity utilisation) could be specified as follows:

1. *Zone of loading of transport unit (i)*
2. *Zone of unloading of transport unit (j)*
3. *Transport mode (m)*
4. *Transport unit (u)*
5. *Weight transported in tonnes (TT)*
6. *Carrying capacity moved in tonnes (TC)*
7. *Loading factor (LF)*

In mathematical notation**:** $LF_{i,\ j,\ m,\ u}$

$$LF_{i,j,m,u} = \frac{TT_{i,j,m,u}}{TC_{i,j,m,u}}$$

Indicator 5

No.	Label	Unit of measurement	Obser-vational unit	Priority level
5.	Annual total number of tonnes transported broken down by size of shipment (weight), value of shipment[2)] (Euro), trip distance (km), containerisation (yes or no) and type of transport unit	Tonnes/year[1)]	O/D pair (main leg)	F

1. Tonne-kilometre figures may be derived when combining tonnes/year and distance between O/D pairs.
2. It might be considered to present the value flows in the same spatial and modal details as volume (tonne) flows.

A database record with these specifications could be represented as follows:

1. *Zone of loading of transport unit (i)*
2. *Zone of unloading of transport unit (j)*
3. *Transport mode (m)*
4. *Transport unit (u)*
5. *Shipment weight class (s)*
6. *Shipment value class (v)*
7. *Trip distance class (d)*
8. *Cargo type (c)*
9. *Weight transported in tonnes (TT)*
10. *Value transported in Euro's (VT)*

In mathematical notation**:** $TT_{i, j, m, u, s, v, d}$ and $VT_{i, j, m, u, s, v, d}$

Indicator 6

No.	Label	Unit of measurement	Observational unit	Priority level
6.	Average number of days of use of the transport unit	Days	Zone	D

This indicator is intended to describe the availability of the transport unit. In combination with other freight transport supply and demand indicators a global figure might be calculated to monitor the capacity situation in the various transport markets.

A database record with these specifications could be represented as follows:

1. *Zone of registration (z)*
2. *Transport mode (m)*
3. *Transport unit (u)*
4. *Carrying capacity class (tc)*
5. *Number of transport units (TU)*
6. *Average weight carrying capacity (ATC)*
7. *Average number of hours of use (H)*

Indicator 7

No.	Label	Unit of measurement	Observational unit	Priority level
7.	Use of EDI, tracing and tracking of shipment	Yes or no	Zone	D

It is hard to find a meaningful interpretation for this INFOSTAT indicator if the observational unit is a zone. From the point of view of freight transport demand it should say something about the use of EDI and tracking and tracing of individual shipments. What zone should be taken for the aggregation of this shipment related information, the origin, the destination or perhaps the transshipment zone? It seems more appropriate to aggregate to an O/D-relation with perhaps further specification of the kind of shipment as for indicator 5. Because this indicator is not seen as fundamental, it will not be treated as a separate indicator, but can be treated as part of indicator 5.

The indicators described above can be considered as belonging to four different groups or levels of information, ranging from the goods flow in transport means (mode and unit), the trip, the load (consignment) and the transport chain. These groups are closely related (linked) to each other.

A scheme for bringing together the various data elements of the ETIS goods transport demand indicators by linking the four levels of information in a policy oriented comprehensive database is presented in table 2-2. The following aspects should be kept in mind before linking transport information:

- A harmonised classification should be used to be able to keep hold of the most detailed information possible, such as information on the load (consignment)
- Individual items should be considered as new variables, such as number of trips, number of consignments, etc.
- Some additional variables should be calculated, e.g. weight x consignment distance for tonne-km, capacity x (loaded and/or empty) trip distance, etc.
- Consistency with classification used in trade databases for building transport chains
- Consistency with classification used in other parts of ETIS (socio-economic data, passenger flows, infrastructure links and nodes, impacts).

The concepts in table 2-2 apply to an ideal situation. Then linkage between the elements, especially between the zones of loading/unloading of the successive trips of a specific load, would be possible at a very disaggregate level, i.e. at the level of each distinct trip made by a specific load.

Table 2-2: ***Four levels of individual linked information on goods transport demand***

Observation level and conceptual symbol			**Structure of the variable**	**Measurement unit**
Mode and unit of transport				
1	Zone of registration	z	NUTS 2/3 zone	
2	transport mode	m	sea, road, inland shipping, rail, pipe, air	
3	transport unit	u	tank, swap body, container, pallet etc.	
4	weight carrying capacity	TC		Tonne
5	volume carrying capacity	MC		m^3, pallets, TEU's, etc.
6	effective hours of use	HC		hours
Trip				
1	link to mode and unit of transport		key	
2	zone of starting of the trip	i	NUTS 2/3 zone	
3	zone of ending of the trip	j	NUTS 2/3 zone	
4	kind of trip	k	loaded, empty, pick up & delivery, etc.	
5	total trip distance	DT		km
Load				
1	link to trip		key	
2	zone of loading	i	NUTS 2/3 zone	
3	zone of unloading	j	NUTS 2/3 zone	
3	commodity group	g	NSTR, HS class	
4	cargo type	c	dry bulk, wet bulk, general cargo, loaded/empty container, etc.	
5	EDI, tracking & tracing	e	yes/no	
6	weight of load	TT		tonne
7	value of load	VT		EURO
8	volume of load	MT		m^3, pallets, TEU's, etc.
9	distance O/D load	DS		km
10	price of the transport of the load			EURO
Transport chain				
1	link to load 1		key 1	
2	link to load 2		key 2	
3	link to load 3		key 3	
..	…….			
n	link to load n		key n	

The element *key* in table 2-2 denotes various sets of information (coefficients, parameters, etc.) enabling us to build complex information chains. To achieve this one must at the bottom have information at a very detailed level such as the single load.

If aggregate information is used, it will be difficult to produce the required linkages, especially from the load to the transport chain. A reverse scheme must, therefore, be taken into consideration as a way to assemble the different elements of the transport chain, going from the identification of the shipment to the characteristics of the transport movements (trip, mode and transport unit attributes). Information consistent with such a scheme could be compiled through shippers' surveys in the short-term and through shippers' or major transport integrators' EDI systems in the medium/long-term. The drawback would be the cost to collect new information, but the advantage would be the assurance from the beginning that different elements could be linked at a very disaggregate level. At the same time one could also collect information about economic and logistics determinants of transport.

A comprehensive review procedure was used to identify the level of compatibility between the actual proposed indicators and the structure of data available from different sources, but fulfilling a similar purpose. The conclusion from the extensive search for available and relevant statistical information was rather simple. Traditional statistical data on European goods transport flows are mostly restricted to a unimodal registration at an annual basis of tonnes and/or tonne-kilometres broken down by commodity group and only specifying country of origin and country of destination (NUTS 0). The lack of more detailed O/D information is seen as one of the major shortcomings of the existing transport statistics. To test out in some detail the possibility of establishing relevant databases by combining and restructuring available data sets (databases), the Transalpine Pilot Study was carried out. This pilot study is described in chapter 4.

2.1.2 Pilot study TRANSITIE

The conclusions and recommendations from TRANSITIE - focused on the creation of a transport chain database - have been incorporated in a detailed project plan for the Dutch central bureau of statistics (CBS) for improving the quality of both the trade and transport statistics and implementing the proposed methodology. To point out the possibility of the proposed methodology, a test has been performed in linking available trade and transport data for one specific mode: maritime transport. Procedures and results are presented in this chapter.

The basis for the new international goods transport statistics will be the transport data of sea transport, inland shipping, road transport, air and pipelines. The starting point is that the transport data should be as complete as possible. This applies mainly to road transport (including foreign transporters), pipelines and transport by air.

Transport chains will have to be established based on the transport data mentioned. This can be done in the following manner:

1. Imports and exports from the Dutch AAD (Aan-, af- en doorvoer) set of data should be linked. All needed transport flows for this linkage should in principle be found in a transport database. The assumption is thus that import and export are integral observations.
2. Information on entrepôt from the Dutch AAD, possibly distinguished by import, export and transshipment should be linked. In principle all relevant transport flows should be traceable, in other words it is assumed that entrepôt also involves an integral observation.
3. The transport volume of intra-EU trade remaining after the two steps mentioned above have been carried out, relates to transshipment, assuming that import and export and entrepôt are integral observations. For the determination (partly) of the other used modes of transport in the international transshipment chains for 1997, one can use the registered transshipment (increased sample survey of one month per quarter); for 1998 only two months will be available (January and April). This transshipment should also be linked to the transport databases. For determination of the transport chain in the inter-European transit a separate methodology will have to be developed. Since transshipment in relation to third countries cannot sufficiently be determined on the basis of available transshipment data, a separate methodology should be developed.

After data have been linked, relevant characteristics from the AAD, such as type of flow (import, export or transit) and statistical value should be processed in the transport database concerned. Linking should be performed with databases which are as detailed as possible. For AAD this means that databases should be used in which, for some transport flows, a number of characteristics of this transport might be unknown. This concerns in particular the container indicator and Dutch transport area of loading or unloading.

The following databases were used for testing:

- Database on Maritime Transport.
- Several databases containing trade by sea were available, and three of them were used for linking. The result presented in table 2-3 is based on a database from CBS with both corrections and estimations included (c). Corrections and estimations have been performed by NEA.

An alternative denoted c yielded the best result. This alternative took into account the direction of trade flows (incoming/unloading or outgoing/loading in the Netherlands), the country of origin/loading or destination/unloading (approximately 200 countries) and the Dutch region of origin/loading or destination/unloading (54 regions can be aggregated to NUTS 3). This alternative resulted in the linking of 81 per cent of the flows, a proportion applying to exports as well as to imports, see table 2-3. It seems that the initial target set at 90 per cent could even be reached if further attention was paid to some imbalances concerning the ports of Amsterdam and Rotterdam, oil imports from Egypt and Saudi-Arabia and the flows with the United Kingdom. The creation of transport chains that is the result of this linking process, is therefore, feasible for maritime transport selected in the Dutch situation.

Table 2-3: ***Total result linking AAD-data sets (alternative c) to sea transport database, in tonnes, 1997***

Starting values				
	Import/export AAD	Transshipment AAD	Total AAD	Sea transport
Import	148 076 255	154 029 075	302 105 330	305 857 171
Export	40 094 786	48 760 002	88 854 788	81 875 880
Total	188 171 041	202 789 077	390 960 118	387 733 051
Total linked				
	Import/export AAD	Transshipment AAD	Total AAD	Sea transport
Import	138 464 278	110 317 087	248 781 365	248 781 365
Export	33 213 266	32 983 471	66 196 737	66 196 737
Total	171 677 545	143 300 557	314 978 102	314 978 102
Remainder				
	Import/export AAD	Transshipment AAD	Total AAD	Sea transport
Import	9 611 977	43 711 988	53 323 965	57 075 806
Export	6 881 520	15 776 531	22 658 051	15 679 143
Total	16 493 497	59 488 520	75 982 016	72 754 949
Per cent linked				
	import/export AAD	transshipment AAD	total AAD	sea transport
Import	94	72	82	81
Export	83	68	74	81
Total	91	71	81	81

Considering the difficulties experienced in the Netherlands, the creation of a pan-European consistent transport chain database will be rather challenging and complicated. It cannot be ruled out that similar data problems also exist in other European countries. However, by comparing data from different sources it is possible to check the quality of the data. A quality check on the transport/trade data is recommended for all countries.

2.2 Passenger transport indicators

2.2.1 ETIS indicators and the need for a consistent database

Table 2-4 shows the INFOSTAT indicators that have been made operational in MESUDEMO. As for goods transport some notation principles are used to specify the passenger indicators:

i = origin zone of a trip
j = destination zone of a trip
k = zone of resident population
l = long-distance
c = trip chain
m = transport mode (e.g. passenger car, bus, coach, railway, air plane, ship)
p = purpose of a trip (e.g. commuting, business, leisure, other)
w = type of car owner (e.g. privately owned, owned by business)
d = distance of trips
t = number of trips

p = number of passenger-kilometres.

The following definitions have been adopted by the MESUDEMO project with relevance for the indicators:

- A journey is the act of travelling from one place to another; it suggests the idea of a somewhat prolonged travel for a specific object, leading a person to pass directly from one point to another.
- Trip is a basic concept comprising the homogenous part of a journey as to journey purpose. So if a person commuting by car stops on the way to the job to do e.g. some shopping, the journey is split into two trips, each with its definite purpose.

If we try to give strict definitions to the first five indicators in table 2-4, we could identify two different viewpoints or perspectives based on a different understanding of the indicators, as follows:

- Defining the variables with reference to the trip production by the resident population of zone *k* from each zone *i*, including the residence zone.
- Defining the variables with consideration to the trip production of zone *i*.

The first alternative that is considered to be most reliable, affordable and viable has been chosen. The elaboration of variables even if it is done in a somewhat different way as compared to the case of goods transport largely benefits from the in depth analysis presented in the previous chapter.

Table 2-4: ***INFOSTAT passenger transport demand indicators***

No	Label	Unit of measurement	Observa-tional unit	Priority level
1	Total annual number of passenger trips generated by zone population broken down by mode or combination of modes and trip purpose	Trips/year	Zone	F
2	Total annual number of passenger-kilometres generated by zone population broken down by mode or combination of modes and trip purpose	Km/year	Zone	F
3	Total annual number of long-distance trips generated by zone population broken down by mode or combination of modes and trip purpose	Trips/year	Zone	F
4	Total annual number of passenger-kilometres in long-distance transport generated by zone population broken down by mode or combination of modes and trip purpose	Km/year	Zone	F
5	Average distance of passenger trips made by zone population broken down by mode or combination of modes and trip purpose	Km/trip	Zone	F
6	Annual distance travelled per inhabitant of zone by mode or combination of modes and trip purpose	Km/person and year	Zone	F
7	Annual mileage per car registered in zone broken down by type of car owner (private, business)	Km/car and year	Zone	F
8	Average vehicle occupancy rate for trips of vehicles registered in zone broken down by vehicle type and trip purpose	Passengers/ vehicle trip	Zone	F
9	Total annual interzonal passenger transport flow by trip purpose, mode (or combination of modes) and type of trip chain	Passenger trips per year	O/D pair	F
10	Annual total of passenger-kilometres occurring on territory of zone broken down by kind of infrastructure used	Passenger-km per year	Zone	D
11	Annual total of passenger-kilometres occurring on territory of zone broken down by trip distance class (type of day, time of day, type of traffic conditions)	Passenger-km per year	Zone	D

In the following listing of indicators two columns called

- related variable description
- variable label

are included in addition to those specified in the INFOSTAT project. Both columns intend to present steps in the process of giving operational content to concepts. The formal symbols used are further elaborated upon at the end of the listing.

Indicator 1

Label for passenger indicator 1	Related variable description	Variable	Unit of measure-ment	Obser-vational unit	Priority level
Total annual number of passenger trips generated by zone population broken down by mode or combination of modes and trip purpose	Number of trips by origin, mode and trip purpose	Trip origin Trip mode(s) Trip purpose	Trips/year	Zone	F

Definition: $T^{k}_{i,m,p}$ number of trips generated by resident population in zone *k* from zone *i*, per mode *m* and purpose *p,* per year

Indicator 2

Label for passenger indicator 2	**Related variable description**	**Variable**	**Unit of measure-ment**	**Obser-vational unit**	**Priority level**
Total annual number of passenger-kilometres generated by zone population broken down by mode or combination of modes and trip purpose	Number of passenger/km by origin, mode and trip purpose	Trip destination Trip distance Trip origin Trip mode(s) Trip purpose	Km/year	Zone	F

Definition: $P^{k}_{i,m,p}$ number of passenger-kilometres generated by resident population in zone *k* from zone *i*, per mode *m* and purpose *p*, per year

Indicator 3

Label for passenger indicator 3	**Related variable description**	**Variable**	**Unit of measure-ment**	**Obser-vational unit**	**Priority level**
Total annual number of long-distance trips generated by zone population broken down by mode or combination of modes and trip purpose	Number of long-distance trips by origin, mode and trip purpose	Long-distance: trip origin trip mode(s) trip purpose	Trips/year	Zone	F

Definition: $T^{l,k}_{i,m,p}$ number of long-distance trips generated by resident population in zone *k* from zone *i*, per mode *m* and purpose *p,* per year

Indicator 4

Label for passenger indicator 4	**Related variable description**	**Variable**	**Unit of measure-ment**	**Obser-vational unit**	**Priority level**
Total annual number of passenger-kilometres in long-distance transport generated by zone population broken down by mode or combination of modes and trip purpose	Number of passengers/km in long-distance trips by origin,mode and trip purpose	Long-distance: destination distance origin mode(s) purpose	Km/year	Zone	F

Definition: $P^{l,k}_{i,m,p}$ number of passenger-kilometres of long distance trips generated by resident population in zone *k* from zone *i*, per mode *m* and purpose *p,* per year

Indicator 5

Label for passenger indicator 5	**Related variable description**	**Variable**	**Unit of measure-ment**	**Obser-vational unit**	**Priority level**
Average distance of passenger trips made by zone population broken down by mode or combination of modes and trip purpose	Average distance of passenger trips by mode(s) and purpose	Trip length (km) Trip length by mode(s) Trip length by purpose	Km/trip	Zone	F

Definition: $D^{k}_{i,m,p}$ average distance of trips generated by resident population in zone *k* from zone *i*, per mode *m* and purpose *p,* per year

Indicator 6

Label for passenger indicator 6	Related variable description	Variable	Unit of measure-ment	Obser-vational unit	Priority level
Annual distance travelled per inhabitant of zone by mode or combination of modes and trip purpose	Total km travelled per inhabitant by origin, mode(s), purpose	Total km travelled Total km travelled by mode(s) Total km travelled by purpose	Km per person and year	Zone	F

Definition: $DKM_{i,m,p}$ total km travelled per inhabitant of zone *i*, per mode *m* and purpose *p,* per year

Indicator 7

Label for passenger indicator 7	Related variable description	Variable	Unit of measure-ment	Obser-vational unit	Priority level
Annual mileage per car registered in zone broken down by type of car owner (private, business)	Annual km per car by zone, by ownership	Annual km/car by area Annual km/car by ownership	Km/car and year	Zone	F

Definition: $DCAR^{w}_{i}$ annual mileage per car registered in zone *i*, per type of car owner *w,* per year

Indicator 8

Label for passenger indicator 8	Related variable description	Variable	Unit of measure-ment	Obser-vational unit	Priority level
Average vehicle occupancy rate for trips of vehicles registered in zone broken down by type of car owner and trip purpose	Average vehicle occupancy rate by zone, by ownership	Average vehicle occupancy by area Average vehicle occupancy by ownership	Passenger s/ vehicle trip	Zone	F

Definition: $OCC^{w}_{i,p}$ average vehicle occupancy rate of vehicles registered in zone *i*, per type of car owner *w* and purpose *p,* per year

Indicator 9

Label for passenger indicator 9	Related variable description	Variable	Unit of measure-ment	Obser-vational unit	Priority level
Total annual interzonal passenger transport flow by trip purpose, mode (or combination of modes) and type of trip chain	O/D flows by trip purpose, mode(s)	O/D flows by mode(s) O/D flows by trip purpose	Passenger trips per year	O/D pair	F

Definition: $T_{i,j,m,p}$ total number of trips generated by origin zone *i* and attracted by destination zone *j*, per type of trip chain, mode *m* and purpose *p*, per year

Indicator 10

Label for passenger indicator 10	Related variable description	Variable	Unit of measure-ment	Obser-vational unit	Priority level
Annual total of passenger-kilometres occurring on territory of zone broken down by kind of infrastructure used	Not relevant	Not relevant	Passenger -km per year	Zone	D

Definition: Not relevant.

Indicator 11

Label for passenger indicator 11	Related variable description	Variable	Unit of measure-ment	Obser-vational unit	Priority level
Annual total of passenger-kilometres occurring on territory of zone broken down by trip distance class (type of day, time of day, type of traffic conditions)	Not relevant	Not relevant	Passenger -km per year	Zone	D

Definition: Not relevant.

Indicators 10 and 11 have been considered not relevant in the context of the origin and destination framework that has been basic to this chapter.

Variables can be sorted in independent and dependent variables, e.g. as follows:

Independent variables

$T^{k}_{i,m,p}$ number of trips generated by resident population in zone *k* from zone *i* per mode *m* and purpose *p,* per year

$T^{l,k}_{i,m,p}$ number of long-distance trips generated by resident population in zone *k* from zone *i* per mode *m* and purpose *p,* per year

$D^{k}_{i,m,p}$ average distance of trips generated by resident population in zone *k* from zone *i* per mode *m* and purpose *p,* per year

$DKM_{i,m,p}$ total km travelled per inhabitant of zone *i* per mode *m* and purpose *p,* per year

$DCAR^{w}_{i}$ annual mileage per car registered in zone *i*, per type of car owner *w,* per year

$OCC^{w}_{i,p}$ average vehicle occupancy rate of vehicles registered in zone *i*, per type of car owner *w* and purpose *p,* per year

$T_{i,j,cm,p}$ total number of trips generated by origin zone *i* and attracted by destination zone *j*, per type of trip chain *c*, mode *m* and purpose *p*, per year, for $i \neq j$

Dependent variables

$P^{k}_{i,m,p}$ number of passenger-kilometres generated by resident population in zone *k* from zone *i* per mode *m* and purpose *p*, per year

$P^{l,k}_{i,m,p}$ number of passenger-kilometres of long distance trips generated by resident population in zone *k* from zone *i* per mode *m* and purpose *p,* per year

Further some relationships between dependent and independent variables can be identified:

Relationships

$$P_{i,m,p}^{k} = T_{i,m,p}^{k} * D_{i,m,p}^{k}$$
$$P_{i,m,p}^{l,k} = T_{i,m,p}^{l,k} * D_{i,m,p}^{l,k}$$

Variable $D^{l,k}_{i,m,p}$ is defined as average distance of long-distance trips generated by resident population in zone *k* from zone *i* per mode *m* and purpose *p,* per year. Values can be calculated from the $D^{k}_{i,m,p,}$ based on the definition of long-distance trip.

The demand both on data and on functionality of the future ETIS strongly depends on the requirement of defining not only a procedure for database creation, but also on a reliable maintenance procedure with flexibility to include new data sets and new layers as needs arise.

2.2.2 Passenger data availability

For passenger transport there is much less relevant, international statistics available than for goods transport. The reason is due to the fact that foreign trade by transport modes has been linked to fiscal policies that have traditionally needed very detailed information on trade aspects.

On the national level availability of passenger transport data relevant for the variables has been checked for countries in the European Economic Area (EEA) as well as for CEEC and some other countries. For each country the following data sources have been checked:

- Specific surveys by transport mode
- National travel surveys (NTS)
- Other specific surveys.

A significant number of European countries have collected valuable data about long-distance passenger flows, including border-crossing traffic, in household and/or passenger surveys, and some very few countries update at regular intervals an interurban O/D matrix, mixing observed and modelled data and including international flows. In some countries there hardly exist any harmonised passenger data on national level. Thus there is quite an effort to be made to enrich the national databases in order to be able to complete national O/D matrices, before international procedures can be drawn up. Considering past experiences, the level of ambition of such a goal must not be underestimated.

Available information on passenger transport demand on the international level is also poor, apart from air and maritime transport (because of their character of closed systems). There are very few O/D related data available in international passenger transport. Some national travel surveys specify O/D by transport modes on a NUTS 0 level for neighbouring countries, i.e. countries with substantial contact between one another if international travels are included in the survey.

The situation regarding the availability of passenger data can be summed up as follows:
- Air passenger data concerning flows between airports are currently available
- Rail data could be created from ticket sales, but the competitive aspects make it difficult to obtain these data
- Maritime data are, despite it is a closed transport system, not easily available

- The most difficult situation is faced concerning road travel, despite the quality of the five years census by UN-ECE. Some kind of *best practise* on the national level may be identified as an example to follow
- Data from the European Travel Monitor are commercially available, but the database is limited to tourist travel. The reliability of the data should be further investigated before conclusions can be drawn as to relevance.

2.2.3 Mirror bias error: An example on France/Italy passenger flows

The analysis intended to test out how available data sources could be used to supply more relevant statistical information and to find out something about the potential for harmonisation of quality, availability and suitability when using national data in two countries. The study was limited both in scope and scale according to available time and costs. Geographically the test was limited to only two countries (France and Italy). The study intended to give the total volume of flows in each direction separately for three transport modes[1] (air, rail and road travel) and three purposes (business, leisure and other purpose). It was also the intention to compile data on accidents, energy consumption and emissions.

It has to be underlined that, at the moment, statistical transport information derived from several different sources, mainly of administrative nature, constitutes an extremely heterogeneous system. In the pilot study it was found that data on French-Italian passenger flows were

- of different statistical quality
- not easily comparable with one another since they refer to data collections that present different definitions, classifications, nomenclature and methodologies.

Despite computing efforts it was not possible to link such figures as presented above to O/D. There were also problems as to data privacy, information reliability and quality of the whole process of data gathering.

The conclusions from the pilot study can be summarised in the following way:

- Total passenger flows between France and Italy could not be described in a way relevant for ETIS with the data made available to the pilot study. At NUTS 0 level it was not possible to find data, which could be used to estimate the transport flow, neither totally nor partially by transport mode
- There were differences in the methodology for data collection, in the data gathering process and even in the definitions of terms
- The problem of data privacy has to be solved
- The idea of quantifying indicators on transport impacts had for lack of available information to be discarded
- European national bodies have to be invited to actively participate in the creation of ETIS because lack of such co-operation may cause the failure of the overall scheme.

[1] Maritime passengers flows between the two countries are limited and for this reason are not taken into account.

2.2.4 Recommendations for passenger transport

The existing situation on passenger data in the different countries makes the path towards ETIS complicated, but not impossible.

A European passenger database within ETIS should include the proposed list of variables and focus on the O/D flows in relation with a zoning system. Data should be provided at NUTS 3 level, since this is the level where transport policy may be best monitored and corridor studies evaluated. From NUTS 3 the data can be aggregated to higher levels and thus be used to feature in global terms the answers to Pan-European policy questions. It is, however, important for TEN-T evaluations, Environment Impact Assessments and for validity and reliability, that data on the national level are collected at more detailed levels i.e. NUTS 4 or NUTS 5.

The existing national data are currently too heterogeneous to be easily harmonised and too scarce to provide a satisfactory European coverage within ETIS. A great effort is consequently needed to get the different countries to enrich and adapt their present system of passenger data collection. The main effort should probably be devoted to the building of a consistent system at the national level, allowing to harmonise the results from passenger surveys within the framework of full scale household surveys, in order to produce reliable estimates of O/D flows. In parallel, the opportunity should be taken to harmonise the definition of concepts and nomenclature of variables within each country and among European countries, in order to reach a consistent set of survey forms at the European level. Communication and co-operation with the national institutions will in this respect be of central importance in order to succeed in building ETIS as a transparent and multi-layered information system.

By using both a top-down and bottom-up approach, and gradually building and supplementing the system with new data collection, coming to a more reliable modelling of O/D matrices based on a better coverage through observed data, a passenger component of ETIS can be developed over time. On the one hand, the result of the household survey planned within the 5th framework DATELINE consortium, benefiting from the previous methodological experience of the EUROSTAT pilot long-distance surveys and of the MEST project, will certainly bring an important contribution to the existence of a European framework for harmonisation. The experience of best practise countries in their combination of compilation and modelling of data to produce O/D matrices should be of great help to establish additional national matrices and combine them at the European level.

Although all this may look as over-ambitious, one should bear in mind that the cost of building ETIS is negligible compared to the cost of inefficient maintenance, expansion and use of the transport system, and the knowledge base it will represent will be of great value. Hence, the building and financing of ETIS should be tied to the extension of TEN-T.

3 TRANSPORT NETWORKS

A transport network is an essential part in ETIS and must necessarily be georeferenced to be fully useful and enter in spatial calculations such as shortest path and nearest neighbourhood calculations. Such a network contains a representation of all the transport modes needed in ETIS (roads, railways, rivers, air corridors etc).

The chapter starts by introducing networks, network concepts and network attributes (chapter 3.1), then follows up with a presentation of principles of ETIS networks (chapter 3.2). Existing representation of networks and their attributes are dealt with in chapter 3.3 and pilots carried out in the context of MESUDEMO are considered in chapter 3.4. Chapter 3.5 draws conclusions and recommendations.

3.1 The networks and ETIS

ETIS is a system intended for analyses on a European scale. This means that it is recommended to define a hierarchical classification system for such infrastructure networks representing Trans-European importance. Such a system should reflect existing hierarchical classification systems already agreed upon in EU, ECMT, UIC, IATA, UN and other international organisations. What concerns the networks of national importance, the national classifications should be used. For the same reason ETIS should also use one European wide zoning system. Using many different zoning systems causes data conversion problems and requires a lot of maintenance. In Europe the NUTS system is defined European wide.

The georeferenced representation of transport networks and their attributes has been made from the point of view of policy relevance. The most important information sources have been:

- Policy documents published by the European Commission on the CTP (Common Transport Policy)
- The INFOSTAT – reports
- Consultations with representatives of other relevant ongoing European RTD
- Literature studies.

Dutch experiences in the field of transport network models indicate that there are substantial advantages of having only one common referential structure to which data can be related. This allows one to see for example if accidents are happening on the same location, so that one can take preventive measures. Thus, some of the variables are location specific and need to be assigned as such. In this way many types of data sets can be related to each other through the spatial linkages and thus enabling complex relational analyses.
The outcome has been identification of three very important concepts: corridor, accessibility and transport chain and a number of attribute groups for transport infrastructure.

Corridors
A corridor is defined as a sequence of strongly interconnected regions. The actual definition of a corridor depends on the focus that in the ETIS context is the CTP. The corridor concept can be used when deciding on future infrastructure investments that need to be viewed in a larger inter-regional or even Pan-European context.

Accessibility
INFOSTAT defines accessibility as a measure of how well regions, cities or other units are interconnected via infrastructure networks. From an infrastructure perspective accessibility can be interpreted as a measure of the centrality of network nodes.

Transport chains
A transport chain consists of a sequence of transport modes used to carry a certain quantity of goods from its original origin to its final destination. Along the chain, one or more transshipments may take place. As such direct flows of goods may be regarded as a transport chain without transshipment.

For transport infrastructure the following groups of attributes should be given special attention.

Construction data
For planning, monitoring and evaluating purposes some information on year of construction and method of financing, including share of EU funding, is essential.

Flow supporting facilities
This group of variables refer to all relevant equipment along links and at nodes that support and guide the flows of transport means, freight and passengers. Indication and the development of the level of service offered in the European networks will thus be apparent. Examples may be: automatic traffic control centres, signalling boxes, service stations, warehouses, parking facilities, etc.

Capacity and design
The capacity variables concern the physical lay-out of the links and nodes. The variable may describe *thickness* (e.g. no. of tracks, lines, lanes available) and *gauge* (such as depth of sea lane, track width, height of bridge/ tunnel) of the link. Maximum traffic loads possible should be included e.g. a Level-of-Service (LoS) indicator for the roads, capacity of airports, warehouses etc.

Maintenance and state of infrastructure
To monitor the effects of the infrastructure use and to project future maintenance spending network, specific indicators are needed. No such indicators have been found.

Restrictions on use
This group of variables refer to the various administrative regulations of the use of the infrastructure and of the services provided with it. Examples can be indicators such as speed limits on traffic, use-only tracks or lanes, opening hours, limits to sizes and weights of vehicles, or draught of vessels.

Accessibility and network connection
From an infrastructure perspective there are two dimensions of accessibility that should be dealt with: *spatial reach* of nodes (describing the extension of the catchment area of a node); and *network connection* (describing the network quality of a node in terms of how well it is connected to the high grade networks, basically the TEN-T of the same or other modes).

Quality of transport services

This group of variablesare closely related to the previous and refer to the quality of the transport linkages between nodes in the network, of the transshipment activities at nodes, and of individual links in the network. Quality is in this context congruent with *time friction* and *reliability*. The indicators should also help identifying infrastructure bottlenecks through comparison with capacity and design indicators.

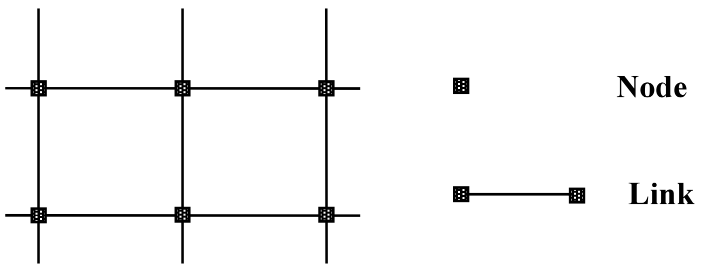
3.2 Fundamental principles for the ETIS network

The following starting points have been defined when designing the structure of the ETIS-network:

- The network should be a potential intermodal network (air, road, rail and waterways).
- The network should include a zoning system.
- Calculations and analyses from the system should be comparable and repeatable.
- The analyses within an ETIS will have different spatial dimensions. Therefore a hierarchical structure of roads, railways etc. should be employed .

Beside those starting points some additional principles for the construction of the ETIS-network have been defined:

- Transportation link as the basic modelling object for the physical infrastructure in the real world.
- The transportation link is the formal representation of (part of) a road, a waterway, an air corridor etc. On a transportation link there is no exchange of traffic allowed and a transportation link must be bounded by exactly two different nodes and it must be possible to travel from each link to each of the other ones. Examples of transportation links are roads, railway lines, air corridors and waterways.
- Transshipment point is the exchange point between the different (transportation) modes. Examples of such transshipment points are railway stations, airports and harbours.
- The mode of transport for goods and passengers cannot change at every place of the network. To be able to change transportation mode certain facilities such as cranes, warehouses, waiting rooms and docking facilities are needed.
- A transportation link should be suited for dynamic segmentation as a long-term option. Dynamic segmentation is a technique to store, query and analyse the attributes of parts of a linear feature without having to modify the whole network feature. The advantage is that the user can define different data sets to different portions of the network elements. When there is no such capability the user should partition lines into segments using (pseudo) nodes. The transportation links must be topologically structured: i.e. they must together form a connected graph (network) as shown in the figure below:
- The zoning system is an integrated part of the network.



Not all traffic and transport statistics can be related to a transport link or a transshipment location. This has sometimes to do with the fact that a particular transport figure is so coarse that it cannot be related to a transport axis but only to a particular relation. Sometime it has to do with the fact that demographic components make part of a transport figure. This means that area features should make part of the ETIS data model.

Selection of network detail
In the GEOSYSTRANS project three hierarchical levels have been proposed in the case of European network application:

International level:	Network of main axes linking major European socio-economic and political places.
National level:	Network of main national axes linking major places of regional importance.
Regional level:	Network of main regional axes linking major places of each country.

Below criteria for the initial selection of networks for ETIS for each mode are suggested.

Roads
The Motorway Working Group of the Transport Infrastructure Committee formulated a *Master Plan for road transport* which came to be the foundation of the technical specifications of Trans-European road network. Their work was the key input to the formulation of general guidelines outlined in the Commission's proposal COM(94)107 in 1994.

The TEN of roads should consist of motorways and other high quality roads that comply with the standards suggested by the working group. The roads must:

- play an important role to long distance traffic; or
- function as bypasses around larger urban centres on certain sections of the net; or
- enable the interconnection of the other transport modes; or
- connect landlocked and peripheral areas with the central areas of the Union.

Railways
The conventional lines of Community interest are defined as those which:

- play an important role for the goods and passenger transport over long distances;
- play an important role in the combined transport operations over long distances;
- enable the interconnection to the other modal networks, and access to regional and local networks.

There are no technical specifications given for the conventional lines, thus, it can be assumed that the international agreement on international standards for the rail network (the AGC agreement set up by UN/ECE) in 1985 is valid.

Air
The TEN of airports should comprise airports of common interest situated within the territory of the Community, and which are open to commercial traffic. The criteria according to which airports of common interest qualify to be part of the TEN depend on the function of the airport. Three types of functionality were identified. To each of these categories the specifications of corresponding airport qualities are described (see L240 24th of August 1992):

- Community connecting points: Key function is to provide links between the Community and the rest of the world;
- Regional connecting points: Key function is to provide links between European countries within EU;
- Accessibility points: Key function is to enhance accessibility for the remote areas to the core of the networks.

Inland waterways

In 1992 the UN/ECMT and the Central Commission for Navigation on the Rhine agreed on replacing the old standards from 1961 with new ones which corresponded better with the traffic flows and which introduced uniform standards for all the European inland waterways (CEMT/CM(92)6/FINAL). Also the concept of secondary networks was introduced referring to networks which link the regional network to the main network or to networks of other modes.

The ports of the inland waterway network function as connecting points between different modal networks of the TEN of goods transport. Thus, to be recognised as ports in the TEN of inland waterways they should be:

- situated on the net;
- functioning as a transshipment centre, thus operating a container terminal as well as roll-on/roll-off facilities;
- open for commercial traffic and not reserved for certain users.

Sea

When the Commission issued the green paper on seaports and sea infrastructures (COM(97)678) the role and in particular the selection criteria for seaports were outlined. The ports were regarded as very important to the TEN-T because:

- they function as connecting points to peripheral areas, or they give access of these areas to the central network structures;
- they play a key role in the efforts to strengthen short sea shipping as an environmentally friendly mode of transport, and as a viable alternative to congested networks;
- they play a central role in the transport system of many of the PHARE and TACIS countries;
- they function as coupling points between land based and sea based transport, and may thus contribute significantly to enhance the efficiency of the overall network;
- they may stimulate growth and trade in neighbouring countries.

The selection criteria include volume criterion for either goods (more than 1 million tonnes of goods per year) or passengers (more than 2 million international passengers per year). An exception is made for ports on the islands in the Aegean and Ionic seas.

The criteria correspond well with the overall political objectives behind the concept of Trans-European transport networks.

3.3 Review of existing networks and attributes

An inventory of a number of important European and national networks for the different modes has been made. In the survey below the summation has been done for networks per mode. Pipelines are excluded because of the private ownership and the specific commodity flows.

In GISCO, only network items describing the segment or node are available. In the extension of the GISCO network made by APUR more information is available such as average daily traffic and accident rate. In national sources, even more information is available, linked to a road segment or a road node. This information concerns for instance traffic flow data and accident information. The goal of the RADEF-project is to connect those national sources in order to exchange and combine them.

The network attributes of GISCO, GISCO-APUR and RADEF have been (globally) compared with the possible attributes from INFOSTAT and other (European) projects. In the Annex the results of this comparison are presented. In the table not all the possible attributes are mentioned, only the most relevant attributes for ETIS have been selected. The other road networks mentioned above are not taken into account either because they only contain very few attributes (NIN, GEOSYSTRANS) or they only are available on a very high level (United Nations).

Infrastructure/ mode	**Networks described**	**Remarks**
Road	RADEF GISCO GISCO-APUR GEOSYSTRANS NIN United Nations Commercial networks (TeleAtlas, Mapinfo, NEA/IWW European networks and AND-mapping)	RADEF is a project from the Western European Road Directors in order to facilitate the exchange of highway information held in various national databases GISCO is the Geographical Information System of the European Commission GISCO-APUR is an update of the GISCO network by the French firm APUR. GEOSYSTRANS is a network made in the Fourth Framework Programme. NIN is the National Infrastructure Network of the Dutch Ministry of Transport, Public Works and Water Management
Rail	GISCO GISCO-APUR GEOSYSTRANS TERN	
Inland waterways	PAWN NEA PC-Navigo	PC-Navigo is a route planner for Western Europe's inland waterways developed and owned by a private Dutch firm. The NEA network is a network used in a production and assignment model on inland navigation predicting future traffic. PAWN is a network used for a model called 'Policy Analyses Waterway management for the Netherlands'
Sea	VESON	The VESON Worldwide Distance Tables is a privately owned

		database with static tables and routing along predefined linking points.
Air	GISCO	
Intermodal	GISCO Inventory of terminals Commercial: Ports database	

Also the rail network attributes of GISCO and GISCO-APUR have been compared with possible attributes from INFOSTAT and other projects. The table, see Annex, does not list all possible attributes, only the most relevant ones for ETIS have been selected. Because of the large number of attributes in GISCO-APUR only a reference to one or more of the extra three subjects is mentioned. The GEOSYSTRANS and TERN-network are left out of this comparison either because they contain few attribute data or they only contain data from a small part of the ETIS-area.

For inland waterways the network attributes of PC Navigo, the Dutch Ministry of Transport/NEA, and PAWN have been (globally) compared with the possible attributes from INFOSTAT and other projects. A major difference between the two sources, PC-Navigo and the Dutch Ministry of Transport/NEA, is that the networks from PC-Navigo contains maximum permissible dimensions of ships to indicate the possibilities of a ship to pass a waterway, whereas the other networks only contains a maximum loading capacity class of a ship. The PAWN network contains the maximum draught of ships, which is an extension of the mentioned networks from the Dutch Ministry of Transport and NEA. However, the PAWN- network is less detailed than the networks from the Dutch Ministry of Transport and NEA. Another difference between the mentioned networks concerns the knowledge of maximum available length, width, height and depth in the PC-Navigo network, whereas in the other networks this information is not available.

For sea the VESON Worldwide Distance Tables provide a global network of ports with distances and routings. The tables include a global network of 1650 ports grouped in three major geographical areas: Atlantic, Indian/Pacific Ocean and Mediterranean. Distances are stored in static tables and the system provides the routing through predefined linking points. Within an area the system provides a single distance given the names of the ports.

The system is expanding by introducing new cluster points in China, Indonesia, India, Africa and South America. Recently 100 cluster ports were added. The company is working on an algorithm, which uses approximately 70 areas and 2000 waypoints to provide the ability to calculate distances virtually from any point to any point.

In short, the information on airports in GISCO is:

- Codes and names for the airport according to different sources (e.g. ICAO and IATA).
- The type of the airport (e.g. civil, civil and military, military).
- Administrative and NUTS region in which the airport is situated.
- Airports code TEN system.
- Airports code TEN type (international connecting points, community connecting points, regional and accessibility points).

In short, the information on ports in GISCO is:

- Codes and names for the port according to different sources (e.g. ISO and UN).
- Administrative and NUTS region in which the port is situated.
- TEN classification of port: end of a link or not.
- TEN classification of port: main inland port.

Ports database contains details of over 4700 ports and terminals and is linked to the distances from the VESON network. Port plans, mooring diagrams, complete specification of port facilities, address and contact details of port service providers. The coupling with the sea network allows us to get an insight in the capacity of the port (draught, mooring capacity, storage capacity).

NEA, INRETS and TFK undertook a study “Inventory of the Intermodal Terminals for Goods and their Characteristics” in 1994. This inventory yielded much information on traffic in terminals. This was a survey held amongst all known major European terminals at that time (779 terminals) to which 367 responded. At present not sufficient and consistent information is available on terminals and their operation.

3.4 Pilots

In this chapter we will deal with 3 pilots. These pilots have as goal to show the comparison between data networks constructed on national level (i.e. NIN and RADEF that have a bottom up methodology) and networks that have a top-down approach (i.e. GISCO and GEOSYSTRANS). The comparison gives knowledge about the detail of the networks and correctness of the network in terms of location as location is an important element for coupling other information to a network.

The following pilots have been carried out:

- The Dutch road network NIN compared with GISCO and with GEOSYSTRANS
- The Dutch rail network
- The Norwegian road network compared with GISCO.

3.4.1 The Dutch road network

By the midst of 1997 the Dutch Ministry designed a completely new network, the National Infrastructure Network (NIN) consisting of nothing else but arcs and nodes, carrying a unique-link-id for internal referencing and for external referencing the least possible number of attributes where added. The aim of the network is to facilitate integration, supporting the further development of a data warehouse, using the clients’ question as leading. At its start the NIN consisted of three infrastructure networks: road, rail and inland waterways all of them using the same data model and existing within one software-environment (ArcInfo®, ORACLE®, SDE®).

Using dedicated maintenance-software the NIN offers high accuracy (< 3 metres), actuality (1 month), interconnectivity between the first three composing networks and an open design towards extensions like Pipelines, Airlines or High Tension lines.

First and foremost the aim of the NIN was to have its data model as “lean” as possible. Using a unique id (the “vak-id”) meant to last during its life cycle, a stable situation has been created. As external keys, road numbering and hectometering for the highways and provincial roads and street names and house numbering (~ranges) for the municipal roads are used. In using these keys the data model can accommodate any "linkage" to either own databases – though in such cases it is more obvious to use the unique vak-id - or databases from other sources or authorities. Using a model such as the NIN outside the Netherlands some difficulties may arise. Referring to the ease in which the Dutch NIN has been modelled using a unique id and road numbers or street names, research has shown that in other Member States the situation is different. It is necessary to have a common understanding and commitment to introduce or adapt the systematic manner of identifying network elements such as roads, streets, roundabouts etc.

The most important lesson to be learned from the Dutch experiences with the NIN is that it is of the highest importance that the approach and systematics of the design and implementation are agreed upon by as many participants as possible. The effort has resulted in:

- the creation of a widely accepted data model throughout the entire field of traffic and transport – and related policy issues ,
- the answer to the need for a generic design which can be used to integrate data,
- the creation of information and presentation of outcome in a geometrically accurate way.

3.4.2 Comparing NIN network with GISCO and GEOSYSTRANS

In the following figures the GEOSYSTRANS and the GISCO networks have been compared with the NIN networks because the accuracy and actuality of the NIN network is known. A minor problem of that comparison is, however, that the NIN does not have a functional diversification. It has a diversification of road ownership and in the Netherlands that is globally comparable with a functional classification. The main roads are owned by the national government, and municipalities own the minor roads.

On a countrywide view the roads (i.e. a view on NUTS 0 regions) of both GEOSYSTRANS and GISCO are comparable with the NIN. The higher the level of the NUTS regions the less accurate are both the GEOSYSTRANS and the GISCO. The figures 3-1 and 3-2 compare results for an area in the western part of the Netherlands which is about the size of a province (NUTS 2 in the Netherlands). Figures 3-3 and 3-4 show the comparison of a part of that province around the city of the Hague which is comparable with NUTS 3 level in the Netherlands. Comparable situations can be found in other parts of the country.

If the situation in the Netherlands is comparable with the rest of Europe this should lead to the conclusion that all network representations can be used for analyses on NUTS 0, NUTS 1 and maybe on NUTS 2. Using GEOSYSTRANS and GISCO on NUTS 3 level is not advisable.

Figure 3-1: Comparison between GISCO and NIN for the province of Zuid Holland (NUTS 2)

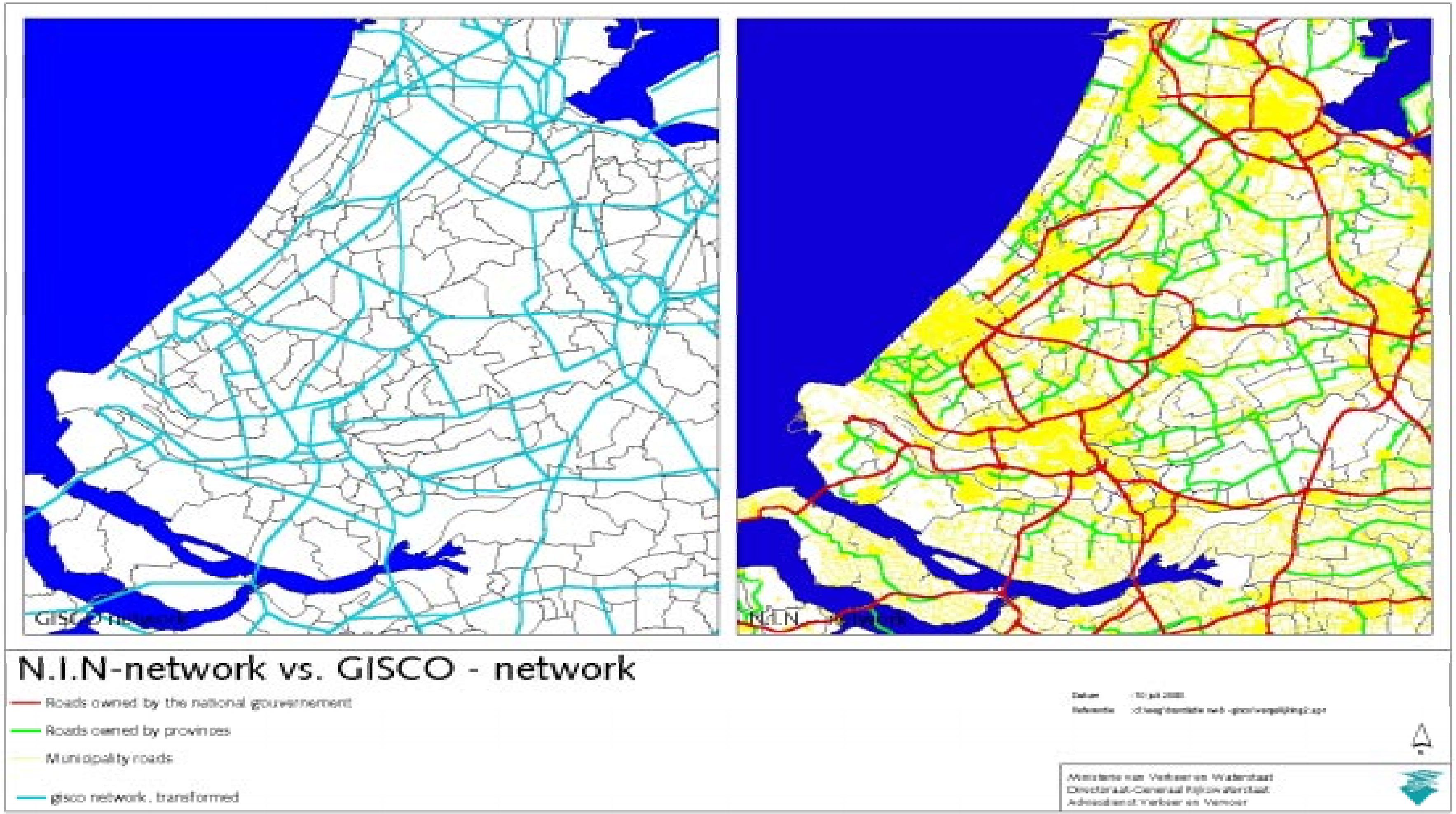


Figure 3-2: ***Comparison between GEOSYSTRANS and NIN for the province of Zuid-Holland (NUTS 2)***

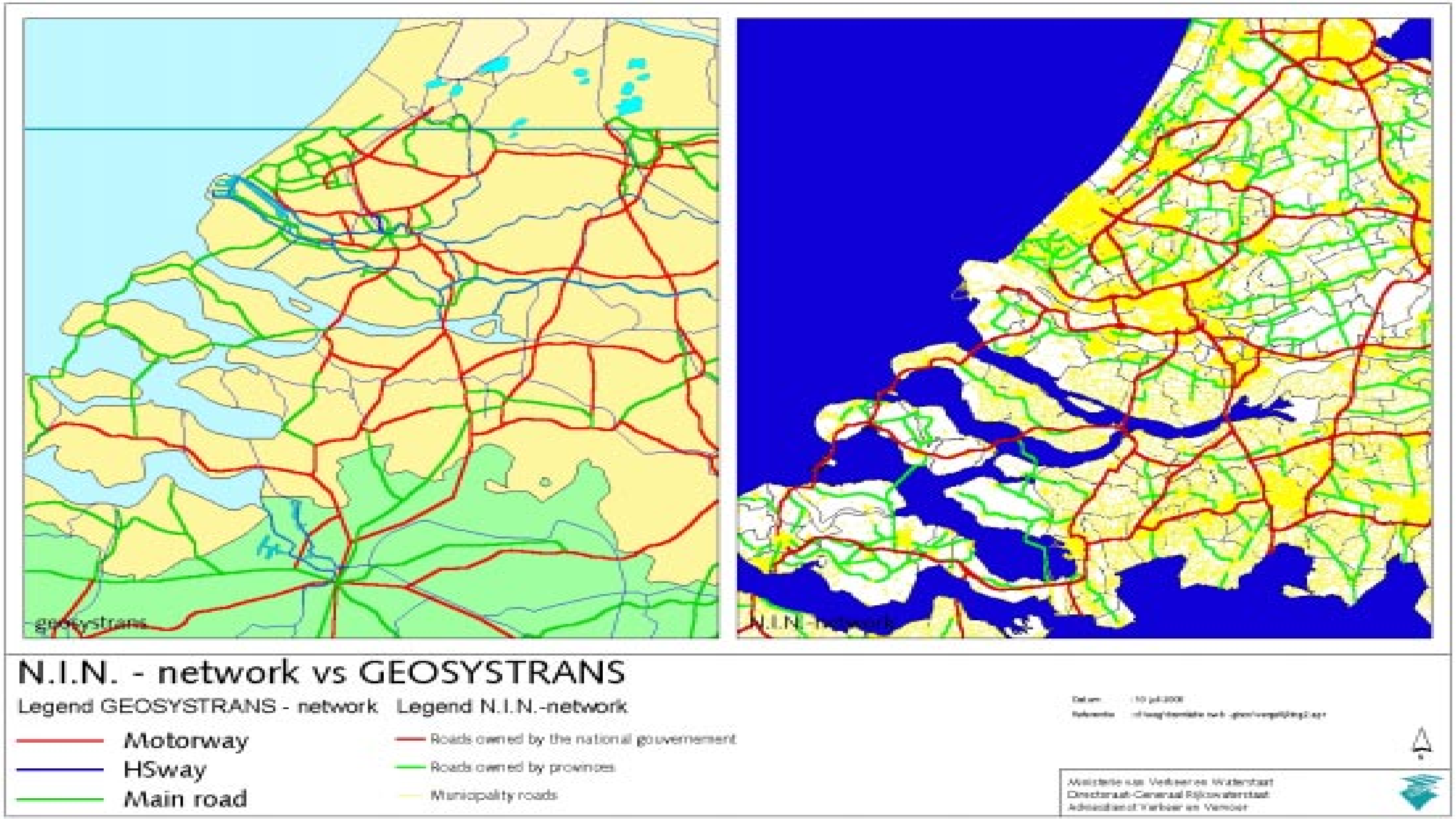


Figure 3-3: ***Comparison between GISCO and NIN for The Hague (NUTS 3)***

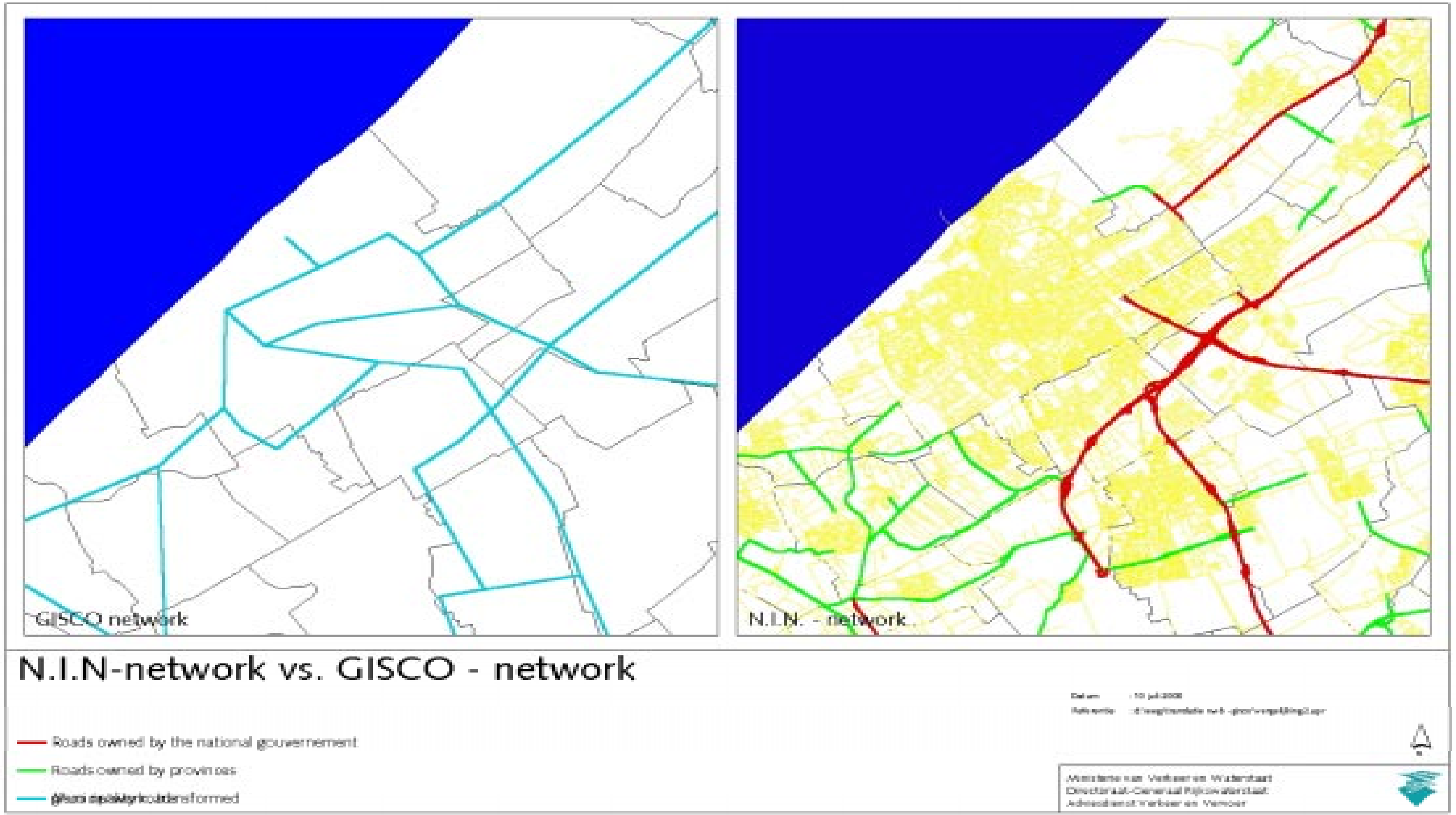
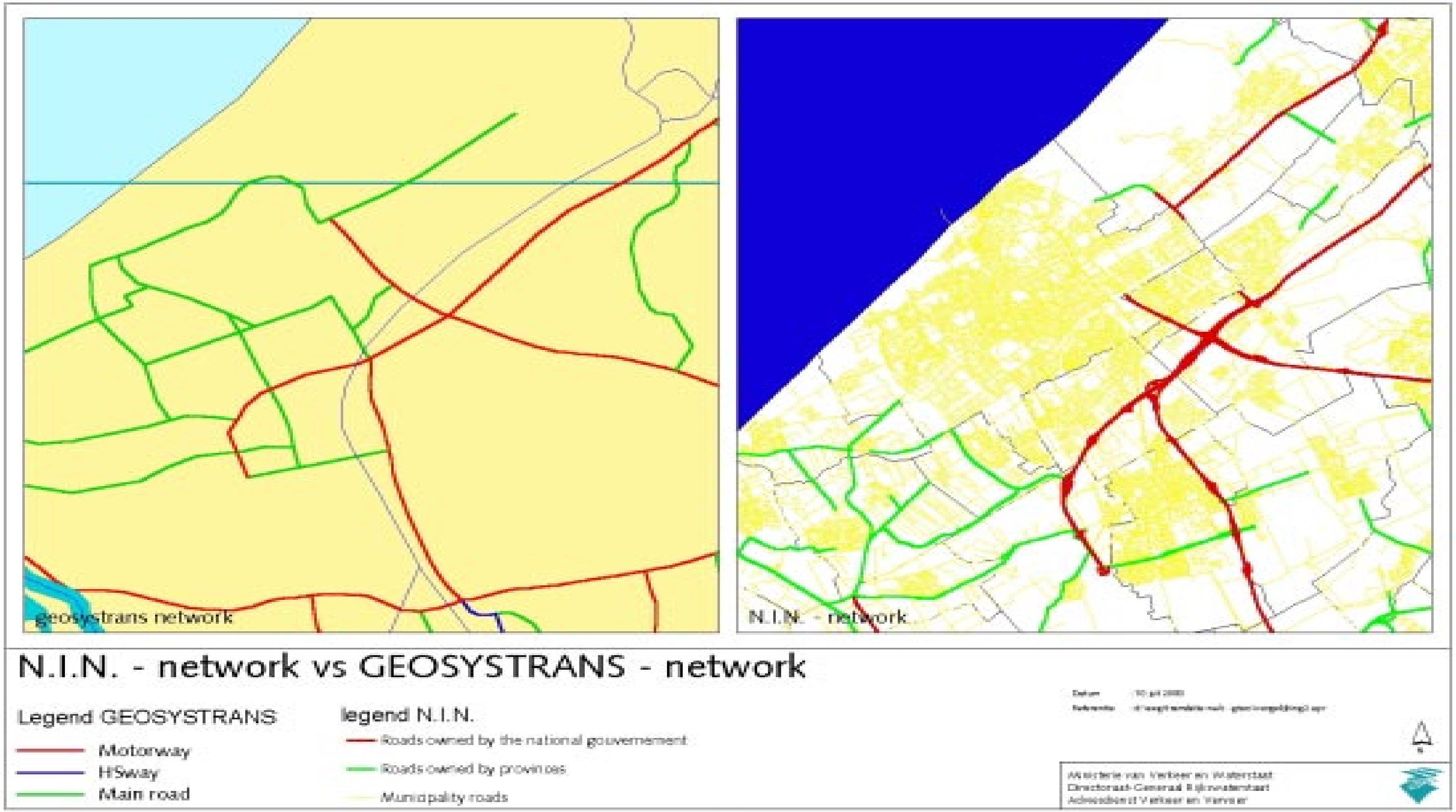


Figure 3-4: ***Comparison between GEOSYSTRANS and NIN for The Hague (NUTS 3)***



3.4.3 The Dutch rail network

The NIN does not only apply for road networks, but for any other inland network as well (road, rail and inland waterways).

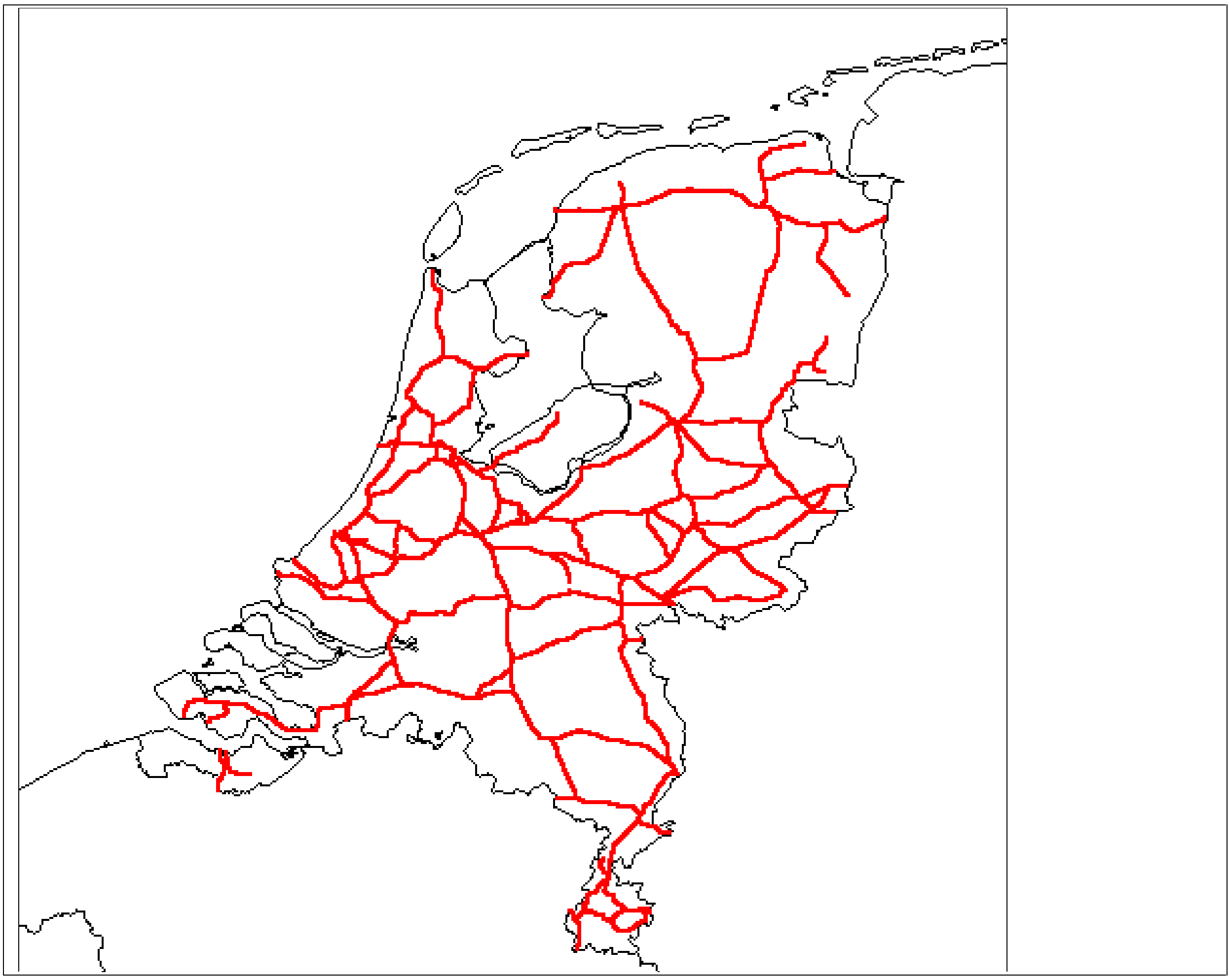
Within the MESUDEMO project a rail-network-data model has been designed based upon the NIN basic scheme. This data model produces transport information – flows, points, areas, quantities, qualities – linked to the network. For transport in particular the design also covers transshipment.

The Dutch rail network covers approx. 5000 km. The information is provided by Dutch Rail (Nederlandse Spoorwegen) and has been translated into the correct data model. The data is matched with existing network information and other "layers" within the NIN. Thus railway crossings are situated where they really are related to the road infrastructure and the "node" is uniquely identified.

Figure 3-5 shows the Dutch NIN rail network and figure 3-6 the GISCO rail network. By comparing both networks only minor differences can be found.

Figure 3-5: ***Dutch NIN – railroad network***

Figure 3-6: ***GISCO rail network of the Netherlands***



3.4.4 The Norwegian road network

In this section the road network as maintained by the Norwegian public road administration is treated. In the figures shown only those Norwegian roads are included that are operated and maintained by the Norwegian public road administration.

The Norwegian road network is compared with the GISCO network. In the figures of the GISCO network the European roads are represented by blue links, the other types of road by red links and the black lines are the borders of NUTS 2 regions.

From the comparison of the European roads (E-roads) of the complete Norwegian road network figure 3-7 shows the whole TERN network of Norway (a selection from the complete network E-roads); figure 3-8 shows the GISCO network. It appears that both networks coincide more or less.

Comparing TERN and GISCO for a selected area in Norway makes possible to study networks in much more detail. It can be concluded that the E-road network is comparable in the two models, but that the other road network is more detailed in the Norwegian network than in the GISCO network.

Figure 3-9 shows the main road network in the counties of Oestfold, Akershus and Oslo. Figure 3-10 is the GISCO network. Figure 3-11 shows the main road network in the county of Akershus; figure 3-12 shows the same area based on the GISCO network.

The coding used in figures 3-9 and 3-11:

1. green colour = European roads (E)
2. red colour = other highways (R)
3. black colour = county municipality roads (F).

Figure 3-7: ***Selection of E-roads (TERN – network of Norway)***

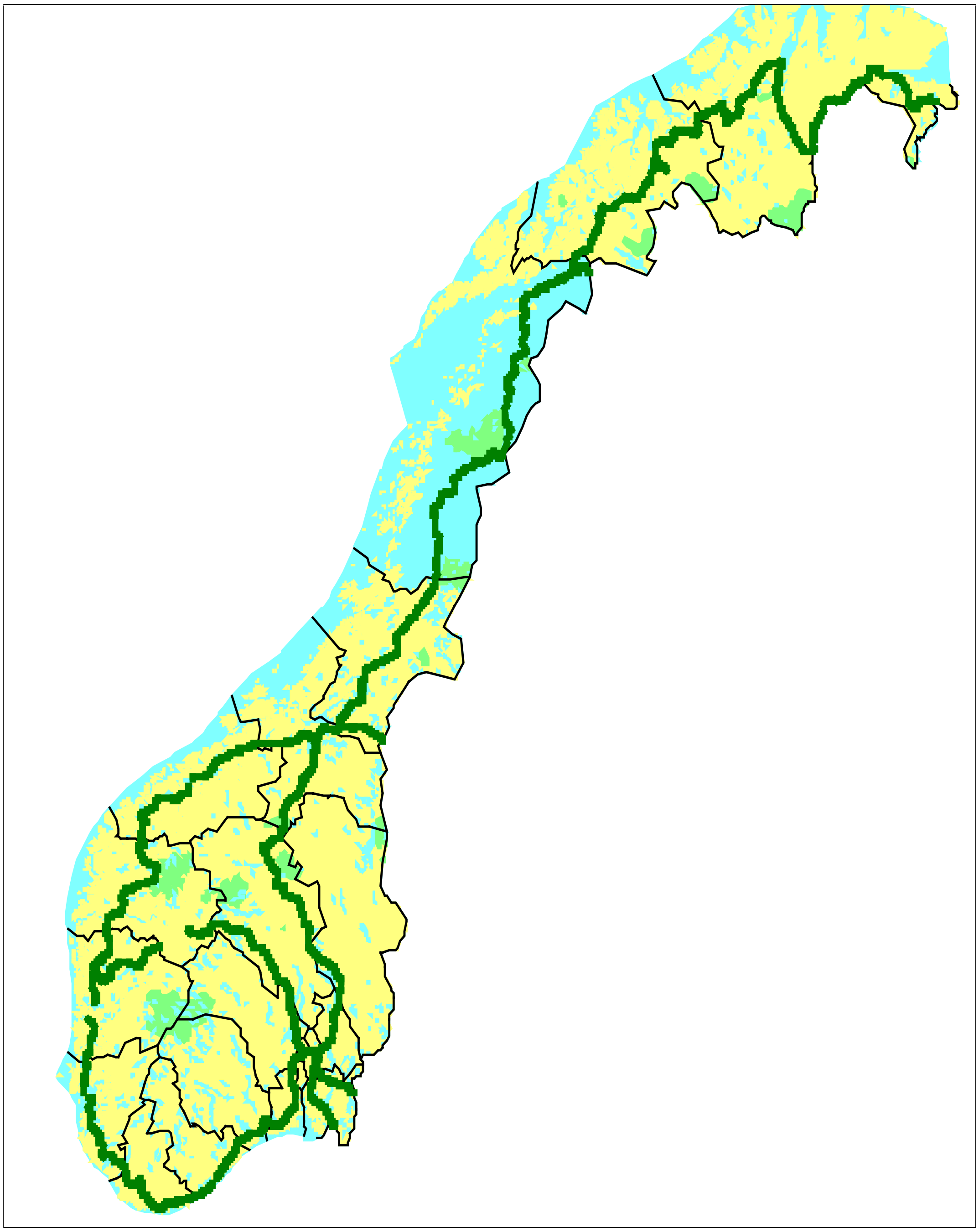


Figure 3-8: ***GISCO – road network of Norway***

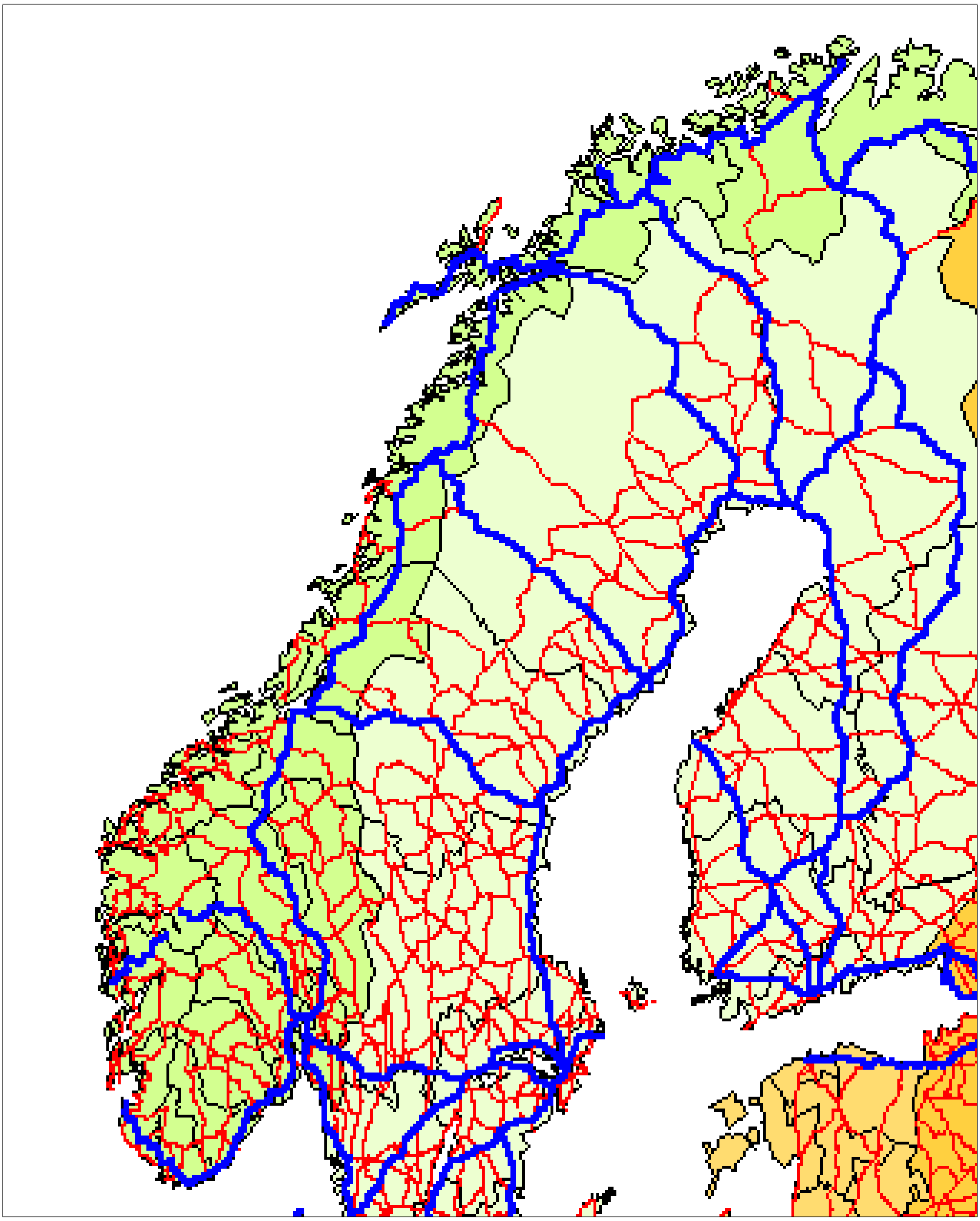


Figure 3-9: ***TERN – Network county Municipality of Oestfold Akershus and Oslo***

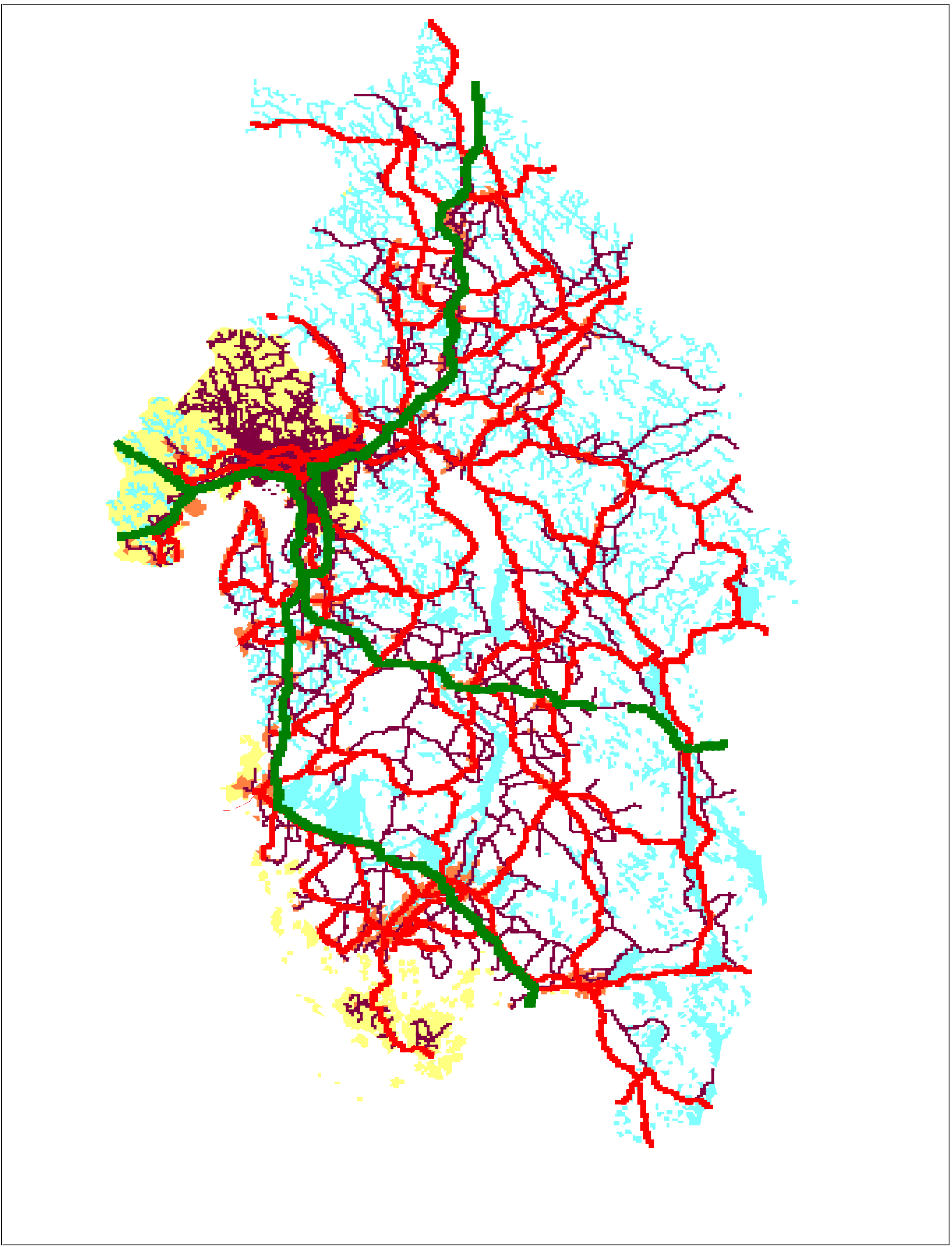


Figure 3-10: ***GISCO Network county municipality of Oestfold, Akershus and Oslo***

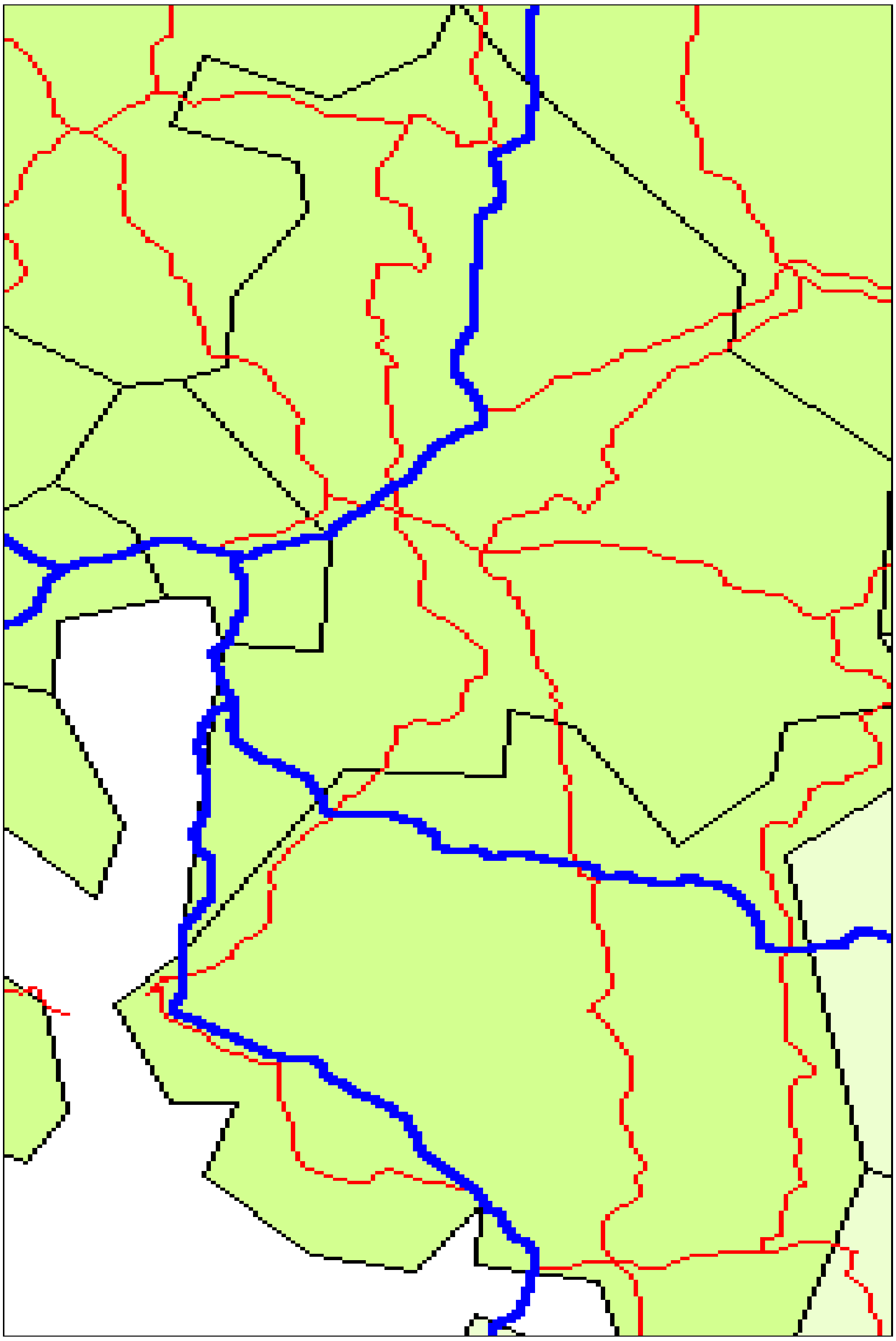


Figure 3-11: ***TERN – network: main roads municipality of Akershus***

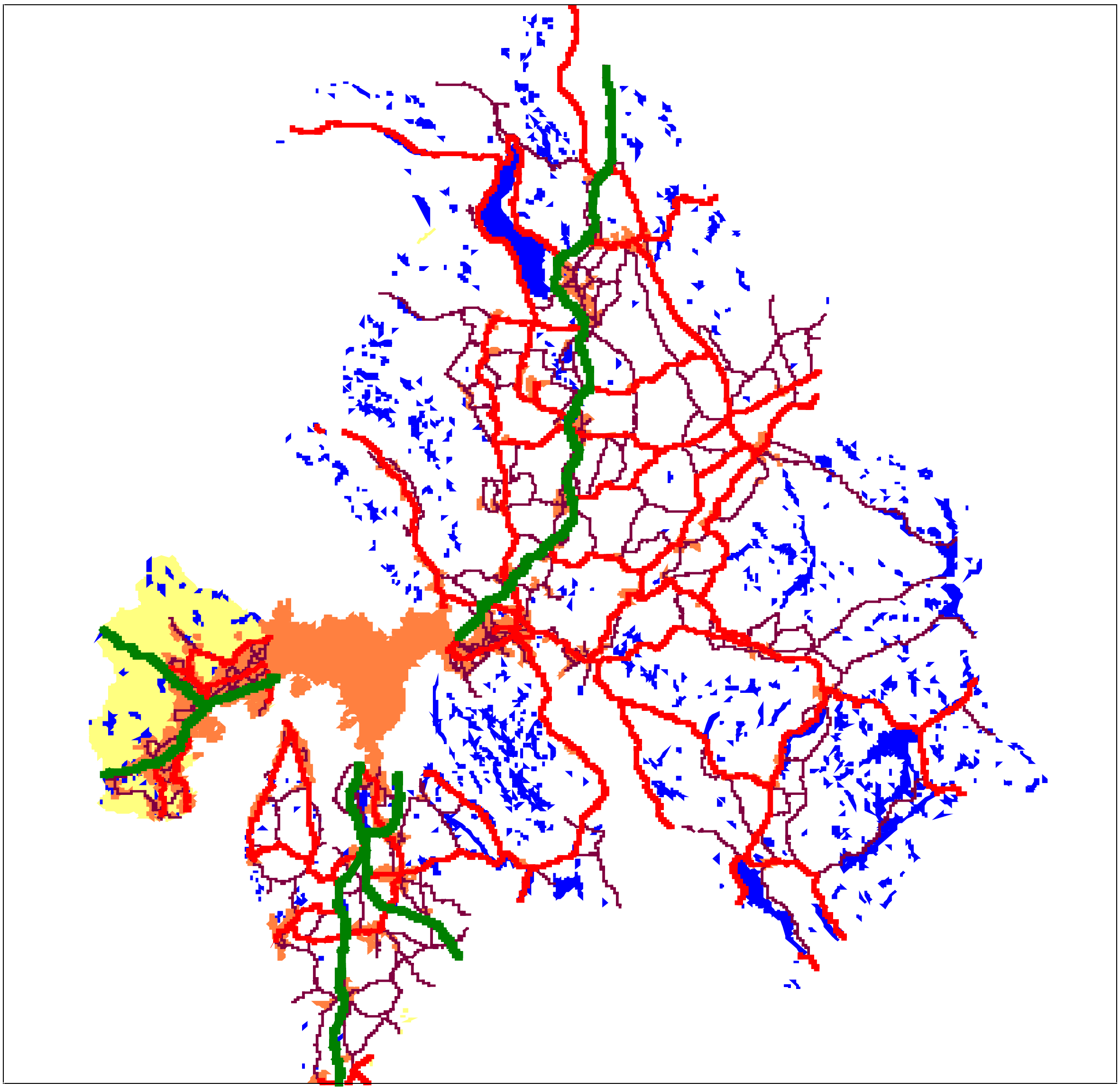
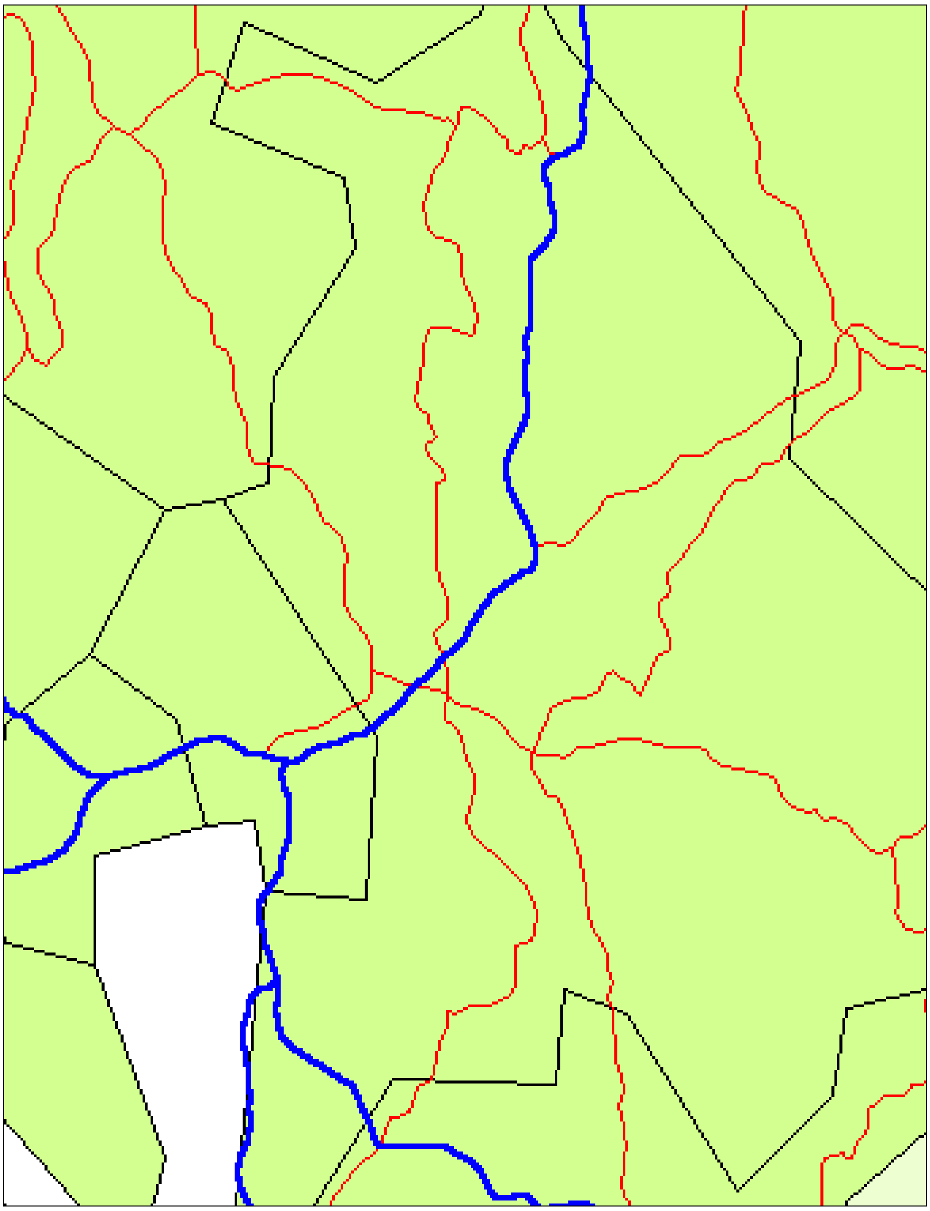


Figure 3-12: ***GISCO – network, roads in the municipality of Akershus***



The pilot shows that road networks supplied by national road administrations are much more detailed than the GISCO and the GEOSYSTRANS networks.. By comparing the national railway network of the Netherlands and the GISCO network of the same area only minor differences were found.

3.5 Conclusions and recommendations

The summary covers two topics:

- The selection of infrastructure networks
- The architecture of infrastructure networks

For the selection of infrastructure networks an extensive empirical research has been carried out. A wide range of infrastructure networks has been studied. For European and national networks an inventory has been made of the underlying sources and the relevant attributes for ETIS. Most commercial networks have not been studied extensively, because the use of these networks is restricted and therefore not suitable for use in ETIS.

The infrastructure networks have been evaluated based on the following criteria:

- bottom up principle: is information received from the countries;
- advanced techniques/architecture: have advanced techniques been used (such as dynamic segmentation);
- advanced attribute handling;
- detail of attributes: is the detail of attributes sufficient;
- restricted/private use: is it a commercial network or can it be obtained free of charge;
- coverage: which parts of the network are included.

One of the weaknesses of the European electronically available data models on infrastructure networks is maintenance. None of the existing data models on European networks are maintained on a regular base. Because of the changes in infrastructure networks the following regular maintenance is recommended:

- road networks at least on a yearly base;
- rail/air networks at least every 5 years;
- networks for other modes at least every ten years.

Road networks

In general, the road network should be able to represent European traffic and transport on a regional basis.

To be guaranteed an up-to-date accurate road network with many attributes one has to use national sources. The national road networks are maintained in most countries on a continuous base. When comparing national and European road networks one of the problems has been that there is no selection process for the ETIS relevant roads. This was also an item at the CONCERTO workshop meeting on (PAN) EUROPEAN INFRASTRUCTURE NETWORKS held in Rotterdam on January 25, 2000. Several representatives of National Road Administrations asked for more guidance from the European Commission in defining relevant roads and attributes.

Generally it is recommended:

- To define the ETIS relevant roads;
- To define the geographical centres of NUTS zones;
- To uniform definitions of attributes;
- To establish an organisation to receive the data from the national organisations on a continuous base.

The RADEF project supplies a network model according to the specifications given above. The RADEF data dictionary provides agreed terms, definitions, categorisations, and relationships for various types of road data, such as traffic details, number of accidents, road geometry and road condition statistics. Common location referencing is a particularly important component. The linking of the national networks into a European network is therefore guaranteed. The data dictionary is the result of extensive co-operative research by the public road administrations. Further guidance from the European Commission to create an ETIS network is, however, necessary.

Networks for other modes

The changes in the infrastructure for non-road networks occur less frequently. The recommendation is to continue the maintenance of existing networks on an ad hoc base using existing European networks:

- rail/air: based on GISCO/GISCO-APUR
- inland navigation: based on the European network of NEA and the Netherlands Ministry of Transport
- sea: VESON.

It is recommended to establish an organisation guaranteeing the quality of the network data. Furthermore a standardisation of the definition of attributes is necessary.

NUTS regions

The classification of Europe into NUTS regions is a classification founded on the number of inhabitants and on the already existing zonal and regional classification systems within each country. Due to the fact that population density varies in Europe, the surface area of the zones differs. A zone on the NUTS 2 level in a relatively highly populated country as the Netherlands or Belgium is much smaller than a zone on the same NUTS level in the northern part of Sweden or Norway. This is not a perfect situation for traffic and transportation studies. It gives extra problems with the selection of the different network elements. Large areas must be handled in a different way than smaller areas of the same NUTS level.

In an ideal zonal system for traffic and transportation studies the size of the different zones do not differ as much as the NUTS regions do but that implies the creation of a new zonal system and the collection of the data needed for that system. This probably has to be considered too much an effort compared with the possible benefits to be gained. The Commission should take up the challenge of having to live with these differences. .

The architecture of infrastructure networks

The architecture of infrastructure networks must be based on common principles used in several projects and networks. The National Infrastructure Network (NIN) of the Netherlands is an example of such architecture. Also other similar ideas like NIN have been proposed, notably in the 4th Framework Programmes Assembling and Bridges. In these projects the GTF (General Transportation data Format developed by Mkmetric), works along the same principles. The detailed structure and the contents of architecture of infrastructure networks are developed in chapter 6.

4 THE TRANSALPINE PILOT STUDY

4.1 Introduction

The main objective of the study was to demonstrate the need and to test the feasibility of building up an ETIS to be applied to a chosen specific area that raises sensitive political issues and where one might expect that the necessary data could be compiled and brought together. The choice fell on ETIS applied to the Alps. Therefore the database has been called the Alpine Transport Policy Information System (ATIS).

In the pilot study the database should contain detailed information on land born vehicle and goods movements for all relevant Alpine crossings, but the study should also implement a transport chain concept and provide information on short sea shipping. No existing data source alone can provide this information. By combining different sources a real added value could be achieved.

The activities in WP 10 follow the structure of the path for constituting ETIS:

- Define the relevant policy issues
- Define the corresponding indicators
- Build up a methodology for computing the policy-relevant indicators from the data
- Build an actual test database using real data
- Calibrate and apply the method to estimate the effect of chosen policies
- Suggest architecture and methodology for generalisation of the ATIS case.

Alpine policy issues are an essential input for the identification of indicators and variables necessary for ATIS. The availability of specific indicators for the region has been analysed from two perspectives: observed and estimated data. The efficiency of ETIS applied to the Alps has been assessed through the interaction of the top-down approach with the bottom-up approach.

To understand the selection of the indicators and variables, the policy issues that have been identified as being of particular relevance to the Alpine region, are listed below:

- Promoting combined transport
- Improving competitiveness of railway companies
- Optimal route selection in the Alpine region (avoiding detoured traffic, justified sharing of burdens of Alpine countries)
- Assessing political measurements and specific regulations (Swiss agreement, Ecopoint system)
- Impacts of toll systems and user fees on route choice in road transport
- Ecological and social impacts of Alpine traffic considering the specific situation of the mountainous landscape
- Assessing the impacts of infrastructure closures (tunnel closures, blockades)
- Assessing the impacts of new transport infrastructures
- Promoting alternatives to Transalpine transport (short sea shipping).

It has to be kept in mind that the variables and indicators proposed for ATIS are specifically related to the more transport related policy questions, which is only one part of the Alpine transport policy dimension. Other parts such as the specific ecological and social situation in the Alps, although being very important, have not been treated in the context of MESUDEMO.

The data architecture of ATIS includes the following modules:

- A supportive module for the data combination procedure
- A module for organisation and preservation of data and knowledge, including a data warehouse and a knowledge base
- A highly organised data core, necessary for properly inter-relating data and feeding the inference engines
- A module for the extraction of information from the data, in order to compute the necessary indicators or perform simulations or predictions, in order to support decision-making
- A set of transport-specific or policy-specific interfaces for the users of the system, including the provision of proper guidance to them
- A set of portals to other systems (organised databases and not data sets), or mediation modules, for proper interfacing to these systems.

The data architecture selected for the pilot study and its concepts are explained in chapter 6.

The chosen issues that the Transalpine Pilot Study can deal with, are the following ones:

A. The pattern of Alpine road traffic
Flow distribution over Brenner
Flow distribution over Mont Blanc
Flow distribution over Gotthard
Pattern of flows from The Netherlands to Italy
Pattern of flows from Rotterdam to Milano.

B. The impact of events and policies on Alpine road traffic flows
Assessment of route changes when closing Brenner
Assessment of route changes when closing Mont Blanc
Impact of restrictive policies on road transport
Assessment of Swiss road toll on road - rail modal split
Assessment of Swiss road toll on the pattern of road traffic crossing the Alps.

Considering the whole range of policy questions identified for the Alpine area and the possible estimates of the Transalpine Pilot Study, the following correlation can be made between the policy questions and the possible estimates:

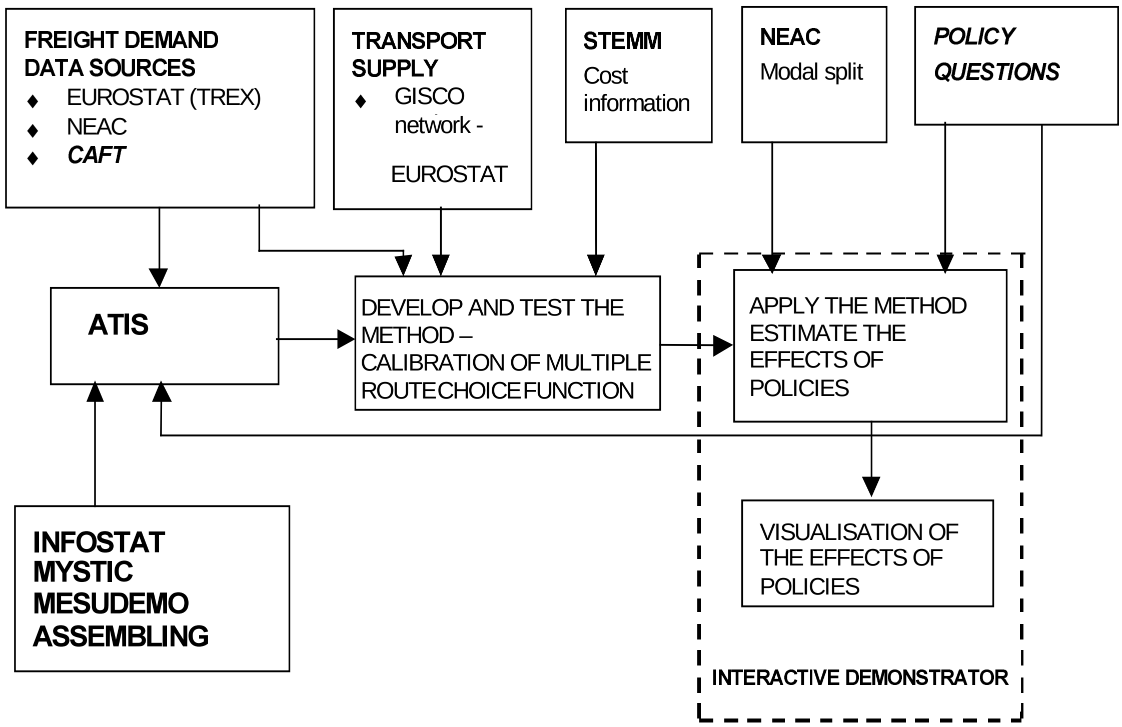
Alpine policy questions	**Transalpine Pilot Study estimates**
Promoting combined transport	Assessment of Swiss road toll (four levels) on road - rail modal split – partial
Improving competitiveness of railway companies	*- not implemented, but it can be done*
Optimal route selection in the Alpine region (avoiding detoured traffic, justified sharing of burdens of Alpine countries)	*- not implemented, but it can be done*
Assessing political measures and specific regulations (Swiss agreement, Ecopoint system)	*- not implemented, but it can be done*
Impacts of toll systems and user fees on route choice in road transport	Assessment of Swiss road toll on the pattern of road traffic crossing the Alps: *four toll levels*
Ecological and social impacts of Alpine traffic considering the specific situation on the mountainous landscape	*- additional instruments have to be implemented*
Assessing the impacts of infrastructure closures (tunnel closures, blockades)	Assessment of route changes when closing Brenner Assessment of route changes when closing Mont Blanc
Assessing the impacts of new transport infrastructure	*- not implemented, but it can be done*
Promoting alternatives to Transalpine transport (short sea shipping)	Flow distribution over Brenner Flow distribution over Mont Blanc Flow distribution over Gotthard Pattern of flows from The Netherlands to Italy Pattern of flows from Rotterdam to Milano

The actual Transalpine Pilot Study could answer to a large range of the policy questions. Some estimates listed above were not computed due to time and resource constraints. Based on the targets to be fulfilled by the Transalpine Pilot Study and on the available information and tools, which were usable within the given constraints, the method of work was designed as shown in figure 4-1 below.

The current approach is based on the following criteria:

- Make use of existing data to the maximum extent
- Make use of existing Pan-European transport models to the maximum extent
- Apply a large-scale information system to estimate the local impact of policies: traffic changes on Alpine crossings.

Figure 4-1: ***Transalpine Pilot Study in ETIS context***



ATIS includes three main parts:

Input data, which in the current application are classified in:

- Data on transport demand, which should be organised in such a way as found in the latest *state-of-the-art* in European research and exemplified in the projects INFOSTAT, MYSTIC and MESUDEMO. The NEAC transport chain database has been made available by NEA for this project together with the CAFT (Cross Alpine traffic database), TREX database from EUROSTAT, national and local (port) databases
- Data on transport supply, represented by the modal and/or intermodal networks. In the present case the application has been limited to the road network, for which the GISCO road network has been made available by EUROSTAT
- Cost information, which consists of a road cost function that is approximated using information from the STEMM project.

Existing models, systems and theories as:

- modal split module of the NEAC transport policy information system
- multiple route assignment technique.

Software and interface tools, as:

- The DELPHI software tool, which includes a multiple route assignment module and manages the interactive demonstrator
- MAPX, which is the GIS visualisation tool.

An *interactive demonstrator* has been built in the Transalpine Pilot Study to demonstrate the possibility of enabling the policy makers to visualise the effects of the applications of different policy options. This interactive demonstrator is able to:

- Show the pattern and structure of existing flows
- Estimate the effect of infrastructure closures on the pattern of road traffic flows
- Estimate the impact of different road tolls in Switzerland on the pattern of road traffic flows.

The interactive demonstrator could be extended in the future in such a way as to be able to extend the range of interactive answers to other policy issues.

4.2 Development of the Alpine freight transport demand database

ATIS freight demand database has to use different sources. The evaluation of these sources and recommendations of the INFOSTAT and the MESUDEMO project imply that the most relevant indicator for goods transport demand is:

No.	Label	Unit of measure-ment	Observa-tional unit	Priority level for ETIS	Priority level for ATIS pilot
1.	Total annual interzonal goods transport flow by commodity group, mode (or combination of modes) and type of transport chain and by ***Alpine crossing***	Tonnes/ year	O/D pair	Funda-mental	

The new dimension that is considered in the Transalpine Pilot Study is the Alpine crossing, which enables the disaggregation of interzonal goods flows by commodity group, mode and type of transport chain over the Alpine crossings. The following sources were used to collect the data to be combined in order to derive the ATIS database:

ATIS: supranational
CAFT 1994 - survey transport data
NEAC 1995 - transport chain data
TRAINS 1995 - transport data
Switzerland 1995, transport

ATIS: transshipment data
SLA Hamburg 1995 – trade / transport
Bremen 1995 – trade / transport
Belgium 1996 - trade
Netherlands 1995 - trade.

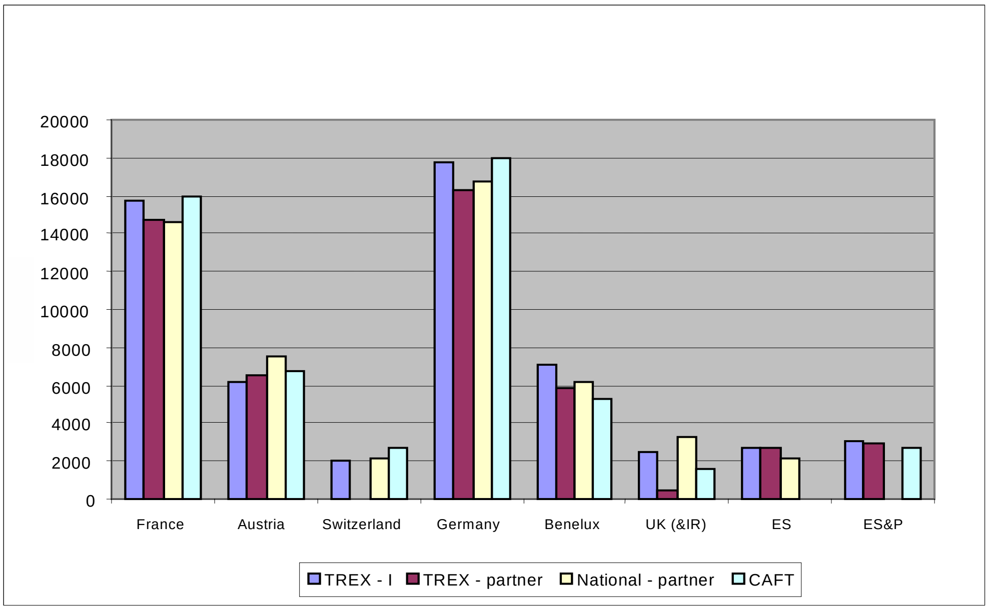
ATIS: national databases
France 1995, transport/trade
Austria1995, transport
TREX 1995 - trade data.
Germany , road 1995, transport 1992-1995
Belgium 1996, trade
The Netherlands 1995, trade
UK 1991-1995 (MDS), trade
Spain 1992-1995 ,trade

The data harmonisation procedure was designed to be simple, robust and understandable and to give reliable results. In order to achieve this goal, the following steps were followed:

STEP 1: Collection and harmonisation of data for comparison
STEP 2: Implementing chain organisation and regionalisation of trade flows at NUTS 2 level by the use of NEAC95 database
STEP 3: Adjustment towards Alpine transport chain organisation
STEP 4: Distribution of transport chain along Alpine crossings

Due to the very different approaches of obtaining the data at the TREX and the Transalpine database, it would not be surprising if significant differences were observed. Bearing in mind the possible bias of enterprise based data collection and customs declaration on the one hand and a sample of lorry driver interviews on the other hand, the rather homogeneous picture shown in figure 4-2 seems almost surprising. The remaining differences can to a large degree be assigned to methodology. Trade flows between Switzerland and Italy or between Austria and Italy do not necessarily cross the Alps, which might explain that the CAFT database shows a lower figure in these cases than TREX.

Figure 4-2: ***Databases comparison - Italian imports by road & rail in Italy, all goods***



CAFT has registered in total 132 million tonnes. Italian international transport flows have the highest share, representing **72.40** per cent of the total. Domestic transport in Austria and Switzerland account for **14.80** per cent of the total and non-Italian international transport accounts for the remainder **12.80** per cent of the total transport volume.

The specific Alpine transport chain organisation and Alpine route related information for a trade volume of ***77 million tonnes*** is included in ATIS database from the total of ***132 million tonnes*** that are registered by the Transalpine database. This is so because only Italian imports/exports in relation with the most important partners have been considered.

ATIS freight transport demand database includes finally the freight flows at the regional level dissagregated per origin, destination, transport mode, transshipment location, commodity group and Alpine crossing.

4.3 Alpine transport supply data for the demonstrator

The Alpine transport supply considered for the pilot study consists of the road network that covers the whole Europe. The GISCO road network was made available by EUROSTAT. GISCO is the Geographical Information System of the European Commission. Typical road attributes are the length of each segment, the road number (national and European), and the type of road segment.

Speed is a very important item not found among the attributes. Knowledge of speed is very important for building the path for each O/D pair. Another important component which is not found among the attributes either is ADT (Annual Average Daily traffic) expressed in PCU's (Passenger Car Units) and/or distributed between cars and trucks.

As the value of average operating speed for freight transport is essential for the Transalpine Pilot Study, speed was estimated using another variable found in the GISCO network, the road segment.

Cost of road transport is expressed as a function of distance and time and is the basic element in determining the optimum route between origin and destination. It is further used as input for the route choice function.

As for assignment, the goods flows have been considered at an aggregated level. It was necessary to find the cost function which makes the average over different types of goods and truck loading. The source for approximation of the cost function is the STEMM project, which makes a review of the VOT (value of time) and fixed costs (per tonne-km).

On the basis of the findings in the STEMM project the following cost function has been approximated (costs in Euro/tonne):

$$ROAD\ COSTS = 0.0458 * DISTANCE + 0.079 * TIME$$

This cost function is further used in the calculation of the generalised transport cost for each O/D pair and for each possible crossing over the Alps. The total volume of goods that is transported by road between an origin and a destination region is distributed over the Alpine crossings by a multiple route choice function, which is calibrated further.

This cost function is used as an average and only for road transport. Future developments of multimodal route choice procedures will have to consider a possible diversity of the cost functions reflecting the differences in transport chain structure (discontinuity of route, type of good, diversity of modes, etc.).

4.4 Development and calibration of the multiple route choice function for assignment

The actual pattern of Italian road traffic flows on Alpine crossings is revealed by the distribution of Italian import and export flows over the crossings. Summarising the flows included in the ATIS database, the distribution of selected Italian imports and exports by road over the Alpine crossings is derived. The results are shown in table 4-1 below.

Table 4-1: Foreign trade of Italy, observed flows (million tonnes) for road transport

Type of flow	Alpine crossing														Total
	P_23	P_21	P_22	P_11	P_12	P_13	P_15	P_31	P_32	P_33	P_34	P_35	P_36	P_37	
Import	3,01	6,38	8,97	0,21	0,02	1,96	0,15	0,18	9,13	0,14	1,55	0,10	0,71	0,23	32,75
Export	2,84	4,47	6,33	0,57	0,15	4,44	0,27	0,26	7,23	0,13	0,49	0,09	0,24	0,70	28,20
Both dir.	5,84	10,84	15,30	0,79	0,17	6,40	0,42	0,44	16,37	0,27	2,04	0,19	0,95	0,93	60,95

The Alpine crossings are described as follows:

P_11 = Gr. St. Bernhard	Switzerland	P_31 = Reschen	Austria
P_12 = Simplon	"	P_32 = Brenner	"
P_13 = Gotthard	"	P_33 = Felbertauern	"
P_15 = San Bernardino	"	P_34 = Tauern	"
P_21 = Mt. Cenis	France	P_35 = Schoberpass	"
P_22 = Mt. Blanc	"	P_36 = Semmering	"
P_23 = Vintimille	"	P_37 = Wechsel (Schäffern)	"

The region-to-region road flows on the Alpine crossings at the level of commodities are further used as observed data in the process of calibration of the multiple route choice function. As the flow from one origin to one destination is distributed over more crossings, the all-or-nothing technique can not be used as an assignment procedure in this exercise. The solution was to calibrate a so-called multiple route choice function, which will enable the simulation of the distribution of flows over Alpine crossings. The logit function was chosen to be the algorithm and calibrated. The calibration consisted in estimating the parameters of the multiple route choice function considering as constraints the (observed) flows over each Alpine crossing, per direction.

4.5 Assessing the existing pattern of Alpine road traffic flows

The structure of selected Italian road freight flows crossing the Brenner, Mont Blanc and Gotthard is given in table 4-2 below.

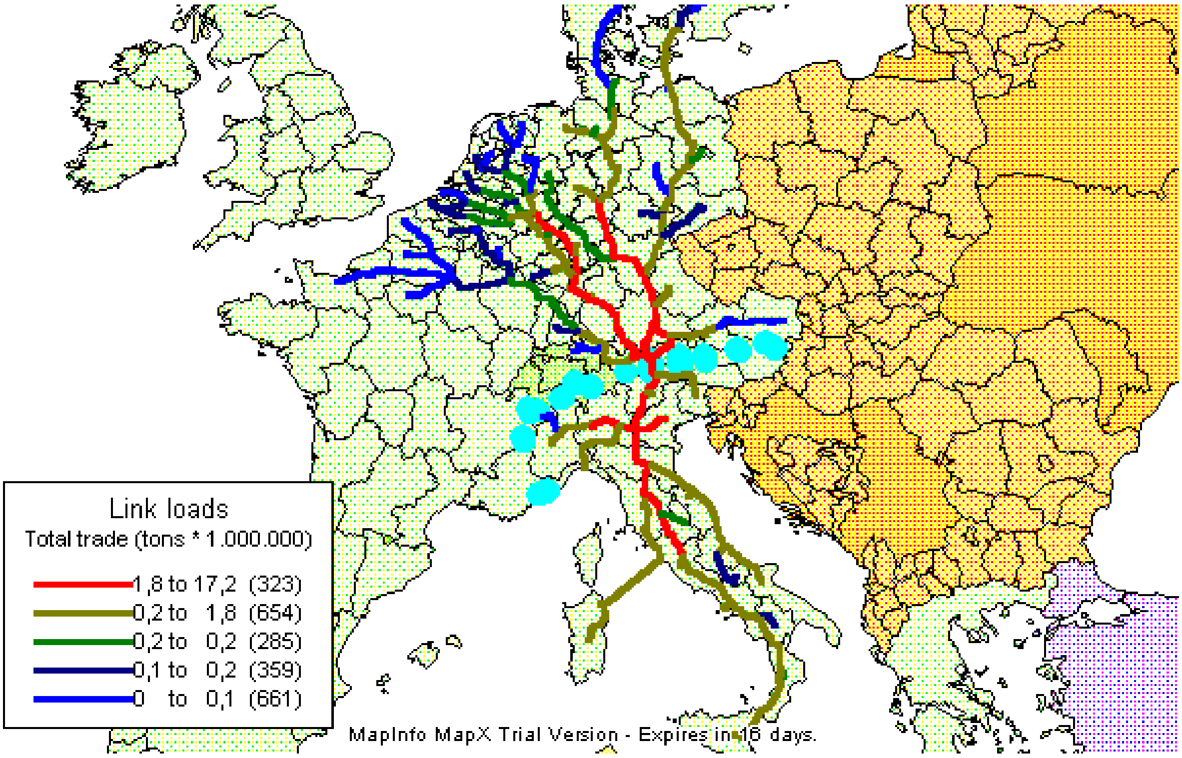
Table 4-2: ***Italian import and export flows by road over Brenner, Mont Blanc and Gotthard – ATIS database. Million tonnes***

Country (partner)	France	BLEU	Nether-lands	Ger-many	Den-mark	Port-ugal	Spain	Norway	Switzer-land	Austria	Total
Road flows over Brenner											
Imports	0.11	0.35	0.34	7.41	0.22	0.00	0.00	0.00	0.00	0.70	9.13
Exports	0.00	0.20	0.17	6.43	0.14	0.00	0.00	0.10	0.01	0.19	7.23
Total	0.11	0.55	0.51	13.84	0.36	0.00	0.00	0.10	0.01	0.88	16.37
Road flows over Mont Blanc											
Imports	4.05	1.71	1.68	1.46	0.04	0.00	0.00	0.00	0.03	0.00	8.97
Exports	3.19	0.87	0.74	1.37	0.04	0.00	0.00	0.03	0.09	0.00	6.33
Total	7.24	2.57	2.42	2.84	0.08	0.00	0.00	0.03	0.12	0.00	15.30
Road flows over Gotthard											
Imports	0.06	0.13	0.12	1.05	0.02	0.00	0.00	0.00	0.59	0.00	1.97
Exports	0.18	0.16	0.13	1.41	0.02	0.00	0.00	0.02	2.52	0.00	4.44
Total	0.24	0.28	0.25	2.46	0.04	0.00	0.00	0.02	3.11	0.00	6.41

The structure of road flows crossing a corridor helps in identifying those flows which could be affected by a change (a restriction) in road transport in that corridor, and by identifying the relations for which an alternative route (i.e. including short sea or inland waterways) could be further analysed.

The structure of selected Italian road freight flows crossing the Brenner is illustrated in figure 4-3 below.

Figure 4-3: ***Structure of road freight flows crossing the Brenner***



The pattern of Italian road freight flows from The Netherlands to Italy, and from Rotterdam to Milano at NUTS 2 regions, is illustrated in the figures 4-4 and 4-5 below. From this picture the distribution of flows over the Alpine crossings could be depicted.

Figure 4-4: ***Pattern of road freight flows from The Netherlands to Italy, by region***

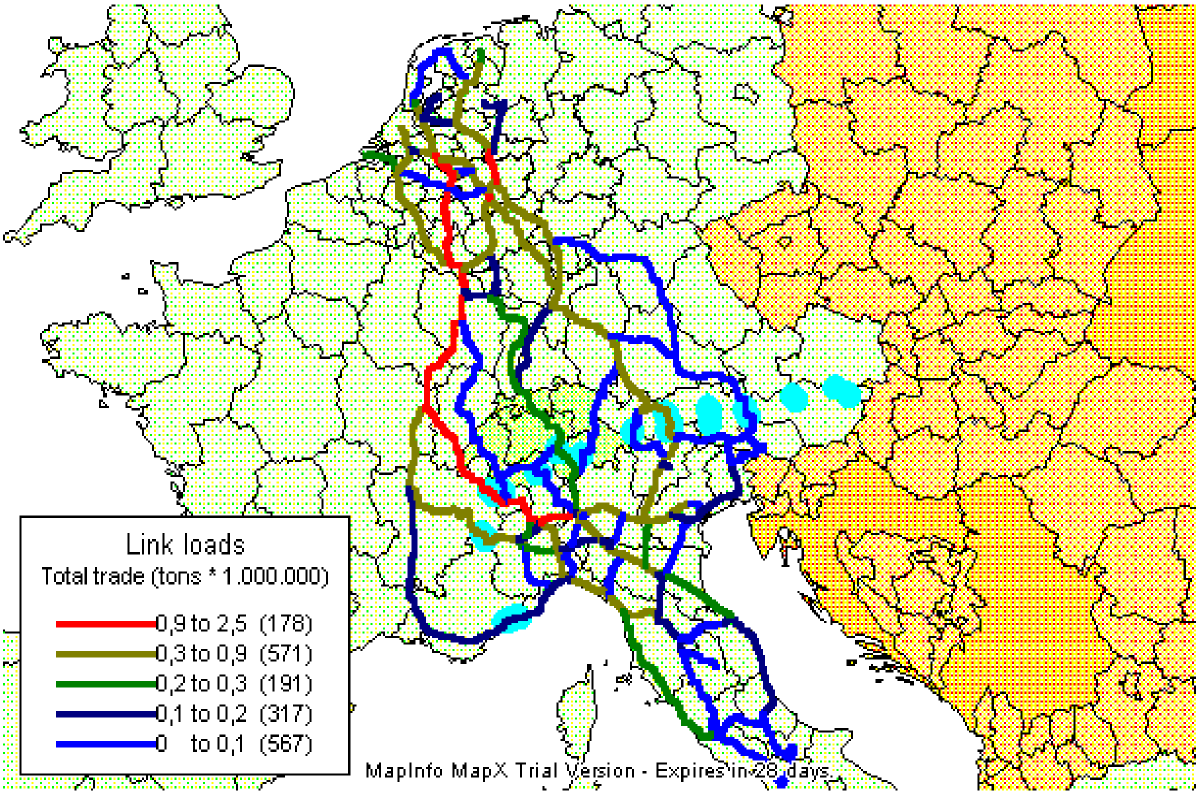
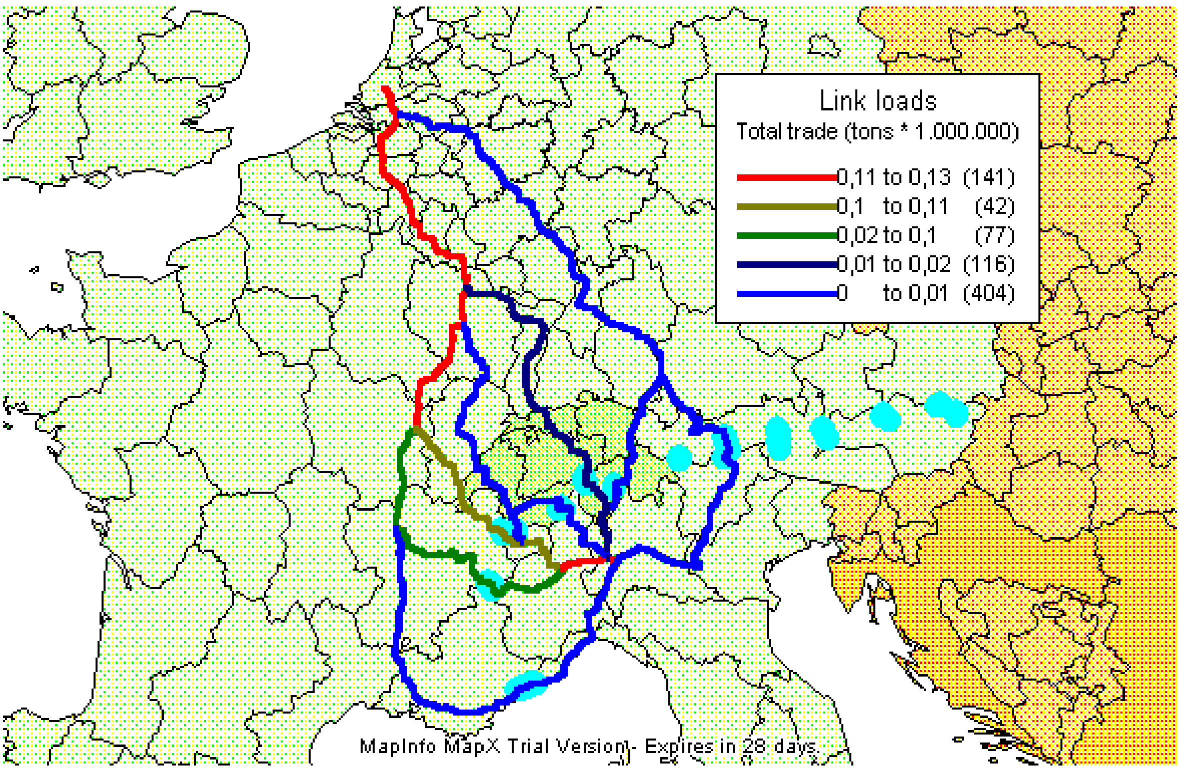


Figure 4-5: ***Pattern of road freight flows from Rotterdam to Milano regions***



A first observation of the pattern of flows from The Netherlands to Italy is that the most used route is the one over the Mont Blanc crossing, with more than 0.9 million tonnes per year. Regarding the transport flows from Rotterdam to Milano (NUTS 2) the most used route is also the Mont Blanc crossing.

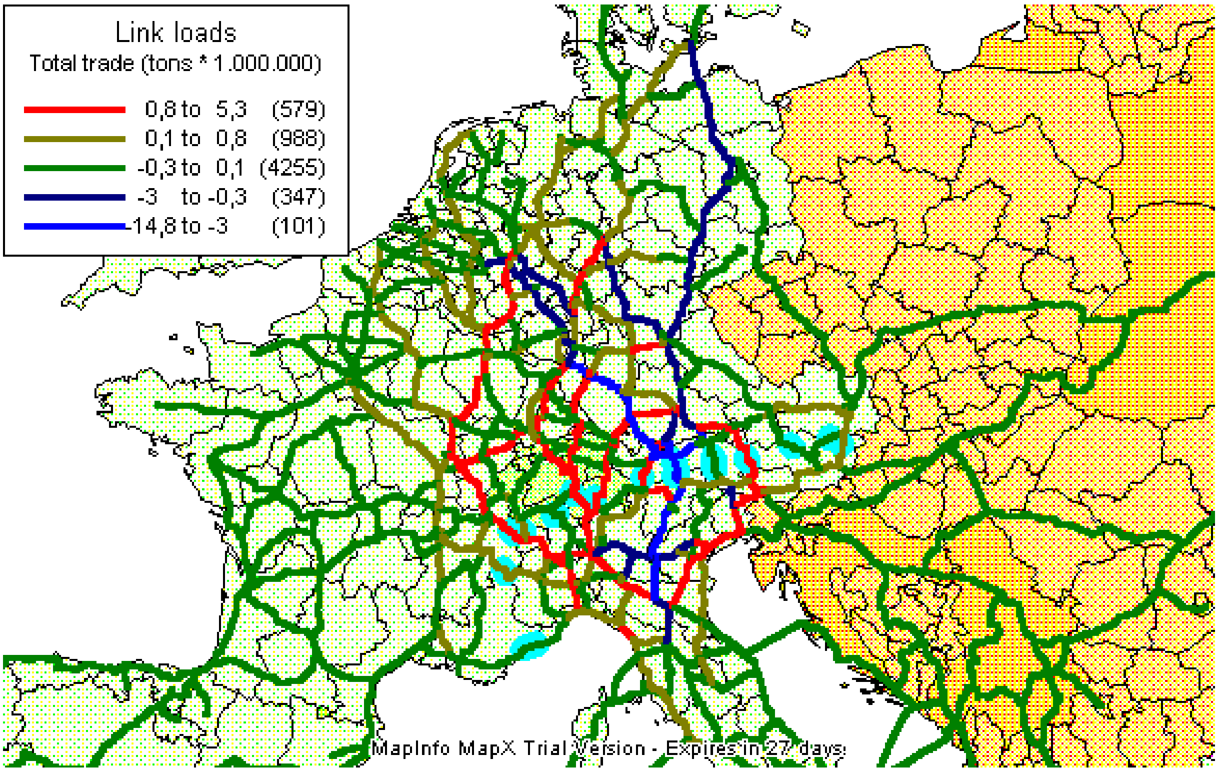
4.6 The impacts of events and policies on Alpine road traffic flows

Infrastructure closures are considered as events and they are simulated within the Alpine transport system by closing the specific links. The events that were considered were the two main Alpine crossings: Brenner and Mont Blanc.

Once the infrastructure closure is implemented in the road network, new road flows are produced assigning the O/D matrix to the network. Road flows that previously took place on the crossing that has been closed, are redistributed over the other crossings.

The effect of closing the Brenner is quantified by the differences in the road traffic flows on each link in the network, as illustrated in figure 4-6 below.

Figure 4-6: ***Change on road freight traffic flows when closing Brenner***



The changes in road traffic flows shown above consider both Italian imports and exports from the selected EU countries. Next, the changes of the distribution of Italian import and export road traffic flows on each Alpine crossing are estimated. The results are shown in table 4-3 below.

Table 4-3: ***Impact of closing Brenner on the distribution of Italian import and export flows by road over the Alpine crossings (new flows/original flows)***

Type of flow	Alpine crossing														Total
	P_23	P_21	P_22	P_11	P_12	P_13	P_15	P_31	P_32	P_33	P_34	P_35	P_36	P_37	
Import	1.02	1.08	1.23	1.04	1.06	2.11	8.37	6.45	0.00	1.44	2.16	1.16	1.22	1.50	0.99
Export	1.00	1.07	1.25	1.03	1.00	1.46	5.69	4.54	0.00	1.31	3.76	1.22	1.15	1.10	1.00
Both dir.	1.01	1.07	1.24	1.03	1.01	1.66	6.87	5.37	0.00	1.43	2.51	1.18	1.15	1.19	0.99

The highest changes in road traffic flows are observed over the Swiss crossing San Bernardino: flows increase 8.37 times for imports and 5.69 times for exports, with the contribution of 13 per cent for import and 14 per cent for exports from initial Brenner flows. Also over the Austrian crossing Reschen the road flows increase 6.45 times for imports and 4.54 times for exports, with the contribution of 9 per cent for imports and 10 per cent for exports from initial Brenner flows.

The differences are relative, as they relate also to the initial flows on the crossings that are affected by the closure of Brenner. As example, Gotthard takes over 22 per cent for imports and 27 per cent for exports of initial Brenner flows, while Tauern takes over 23 per cent for imports and 20 per cent for exports.

As an example of the impacts of restrictive policies on road transport, an assessment of Swiss road toll on road/rail modal split has been carried out. To have a better image of the effect of the level of road toll in Switzerland on the road/rail modal split, four road toll levels were considered:

Level 1: Road toll for crossing Switzerland is 2.00 EURO per tonne;
Level 2: Road toll for crossing Switzerland is 5.00 EURO per tonne;
Level 3: Road toll for crossing Switzerland is 10.00 EURO per tonne;
Level 4: Road toll for crossing Switzerland is 15.00 EURO per tonne.

The NEAC modal split model is used to estimate the new modal split between road and rail. The input to the NEAC model consists of the new difference between the road and rail generalised cost. As toll is applied to the road transport, there are no cost changes for rail transport. NEAC modal split model is applied four times, once for each road toll level. The output of the modal split model is shown below.

EFFECT OF SWISS ROAD TOLL ON MODAL SPLIT

Road toll in EURO/tonne	Modal shift road	Modal shift rail
2	0.976	1.099
5	0.940	1.240
10	0.886	1.460
15	0.832	1.680

As the rail flow volumes are lower than the road flow volumes the relative change is higher for rail than for road transport. After the modal split has been applied, new road flows have been estimated in the O/D matrix of direct road transport for each level of toll. Further, a new assignment takes place and the difference between the new and previous (original) loaded network has been calculated. The impact of Swiss road tolls of 2 and 15 EURO per tonne is shown in figures 4-7 and 4-8 below.

Figure 4-7: ***Swiss road toll of 2 EURO per tonne***

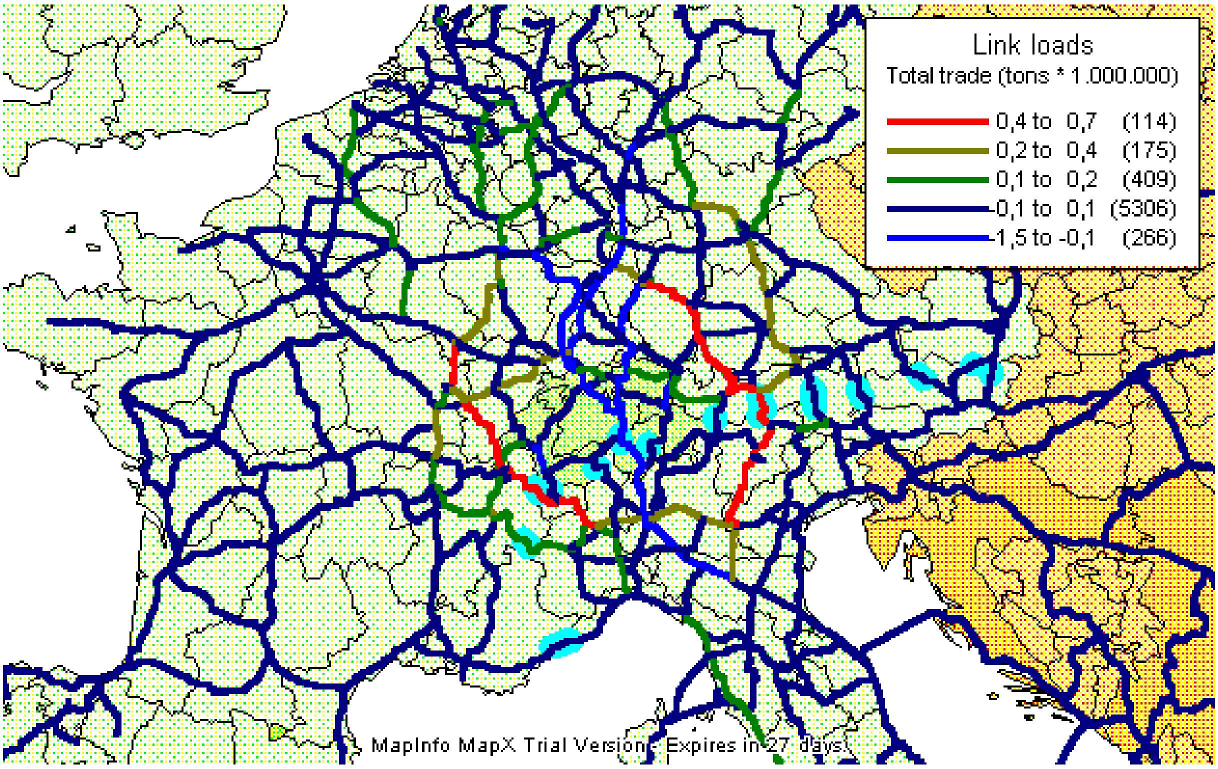
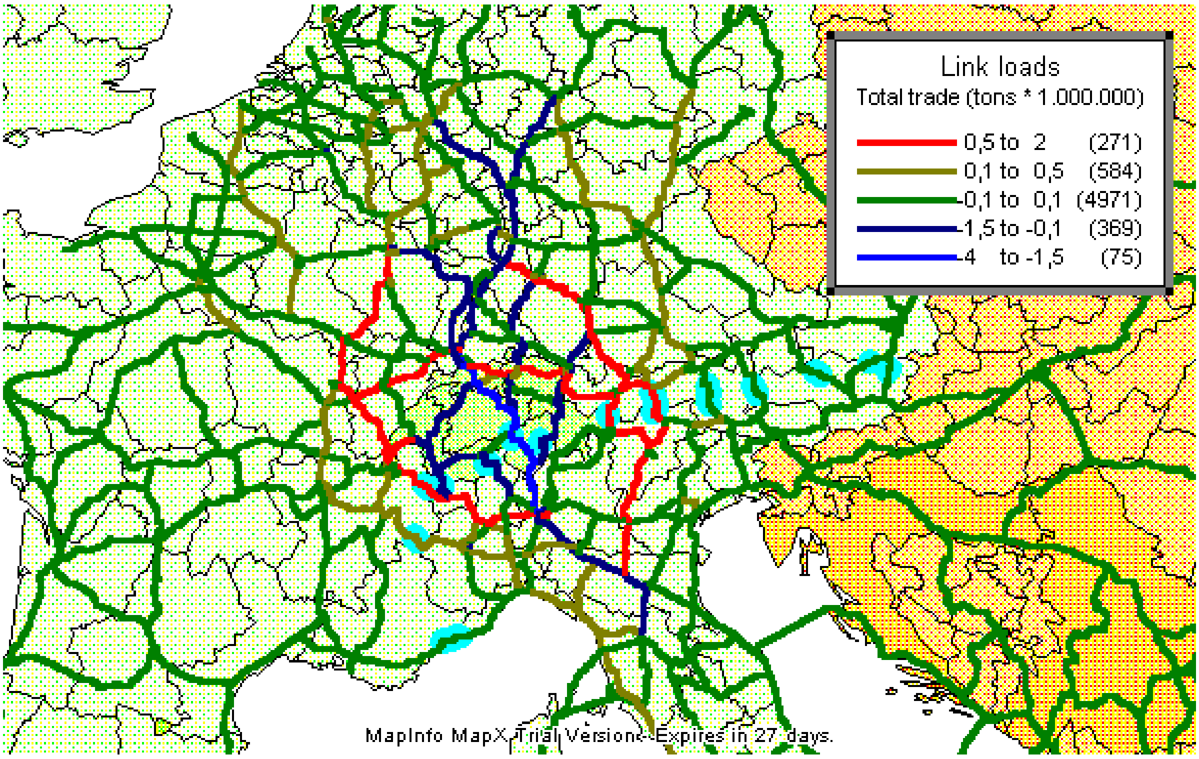


Figure 4-8: ***Swiss road toll of 15 EURO per tonne***



In the case of gradual road toll increase in Switzerland it is observed a tendency of gradual increase of the traffic flows on the French and Austrian Alpine crossings, and also on segments of road situated outside the Alpine area. The increase in road traffic flows over the crossings does not consider the capacity restraints on those links.

4.7 Recommendations for the path to ETIS

The main achievement of the harmonisation process is the successful completion of combining trade and transport data from various sources for Alpine pilot exercise, going through the following steps:

- Regionalisation (NUTS 2) of the trade flows between countries per transport mode and commodity group;
- Implementation of the main transport chain organisation structures in trade flows;
- Implementation of the specific Alpine transport chain organisation in trade flows over the Alps;
- Distribution of trade flows along Alpine crossings over all transport chain organisation structures;
- Assessing impacts related to Alpine transport policy issues using the ETIS approach.

In the Alpine case a rather satisfactory situation was found concerning data on goods transport demand. But before extrapolating this positive view to a future ETIS it has to be kept in mind that the data situation in the Alpine region cannot be compared with the rest of Europe. Due to the focus of public interest in transport through the Alps many joint efforts have been made in monitoring these flows. One main effort has been the 1994 database built on specific surveys on the Alpine crossings of transport flows, this effort being repeated in 1999. It is obvious that, even if the data of the 1994 survey has become one main element of the ATIS database, the recommendation for ETIS cannot be to enlarge the survey to total Europe. But there are some areas and corridors, where comparable surveys, done in a harmonised way, would be very essential for ETIS. Other possible corridors, where the transport flows are concentrating on a few routes and the flows are strongly interregional, concern areas as the Pyrenees (which has been surveyed by France in 1999 together with the Transalpine flows), the Channel, the connections to the northern European countries and the boarder lines to the Central and Eastern European countries. Having in mind that in some of these areas national efforts of data collection have been made, the proposal of harmonising and enlarging them on a European level looks quite reasonable. Such regular harmonised surveys would be very helpful also for checking the quality of the data collected by the regulation for the statistics on freight transport on road.

A major problem faced in the Transalpine Pilot Study concerns data confidentiality. Many transport operators, mainly railway companies but also terminal operators are not willing to hand out their data in a sufficient degree of detail. Here solutions on European level are necessary. The regulation on the statistics of railway transport, which is under development by EUROSTAT just now, points in the right direction.

The confidentiality problem was solved in the pilot study, because existing databases (CAFT, NEAC, etc.) either have been made accessible for the purpose of the project, or were already available at the institutes involved in the pilot study. This problem requires in-depth attention in the architecture. Such a favourable context may not be reproduced for other corridors.

To draw a realistic picture on the important market of intermodal transport and to realise a transport chains approach within goods transport statistics, a specific effort of data collection should be undertaken. Although the access to the part of the information carried out by EDI systems relevant for statistical purposes may be a solution in the long-term, this effort has to be concentrated in the near future on shippers' surveys. As a complement, it would be worth getting some specific information from terminals and harbours. To enable the collection of this information on a European scale, additional provisions are necessary.

The realisation of ETIS revealed to be a complex project. The pilot study in the Alpine area was very useful to understand the process and to fix the steps of the architecture of a future ETIS. Particularly the possibility to build up a European database has been highlighted and is an important added value of the work. Of course this would not have been possible without the use of an existing database and the experience of experts and research offices. Implementing such a process from *scratch* is evaluated to more than 2 million EURO.

The final lesson is that the architecture proposed for ETIS has revealed to be feasible, and the results of the experience in the Transalpine Pilot Study constitute a good basis for demonstrating the various concepts and steps of the architecture. From the experience built up in this pilot process, one would recommend to build up an architecture *by default* to begin with as most of the options are fixed, and to extend it in order to introduce more flexibility in the measurement and analysis of policy impacts.

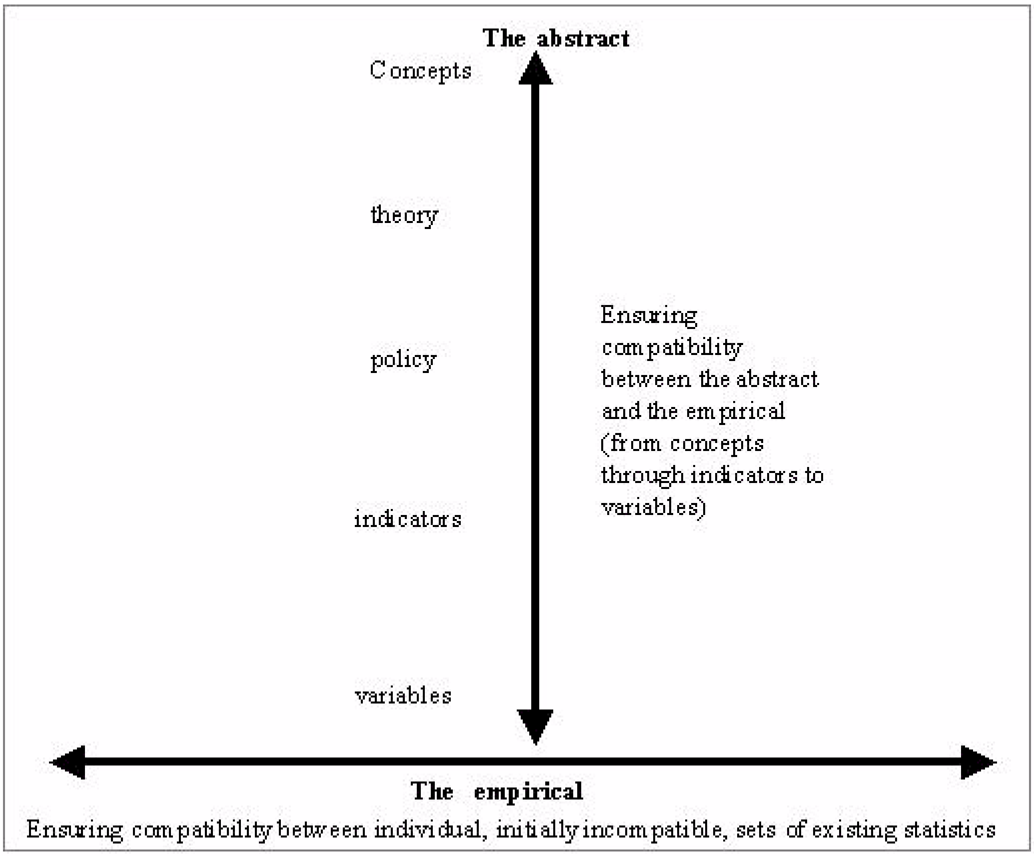
Last but not least, what makes ATIS pilot/ETIS architecture so innovative and different from other products, is the fact that the accent is not specifically on the results (although those are quite promising) of testing policies, but on the path from transport related indicators to policy assessment and impact appraisal by applying all the basic principles on a European level which are recommended for an actual ETIS, with a correct level of detail and in-depth of the indicators.

5 COMPATIBILITY - IMPLICATIONS FOR MESUDEMO

When the objective of MESUDEMO was formulated, it was presumed that methodologies for integrating supply and demand databases would be developed in MESUDEMO, leaving the final stages “only” with the challenge of ensuring compatibility between the two sets of databases. However, for MESUDEMO other compatibility issues surfaced, as difficult and as important as the one between supply and demand.

The compatibility challenge should be viewed along two dimensions, one between already existing data and systems, and the other between how we conceptualise mobility and the data we rely upon to describe and explain it. The first challenge is improving transport statistics and the way statistics are used for guiding decisions on the European level. Along this first dimension, the task is to harmonise data on the European level. Along the second dimension, ETIS must be constructed in a way that makes it possible to link transport effects to their causes. These two challenges are visualised in figure 5-1. It demonstrates two compatibility-dimensions, a vertical one between the abstract and the concrete, and the horizontal between existing statistics.

Figure 5-1: ***Approaches to compatibility***



This chapter is divided in three main parts, the first one explaining why we strive for compatibility between supply and demand for transport in the first place (5.1), the second part dealing with the horizontal dimension in figure 5-1, describing two models for ensuring compatibility between existing data and systems (5.2), and finally a discussion on the challenge of ensuring compatibility between concepts and data (5.3), the vertical dimension.

5.1 Compatibility between supply and demand

The main rationale behind the goal of securing compatibility between supply and demand methodology, is that in order to avoid over- or under-investment in infrastructure, the actual degree of infrastructure utilisation must be closely monitored. This means that information on infrastructure availability and accessibility (supply) must be compatible with information on transport activities (demand), so that both infrastructure utilisation and the forces behind it are made apparent and integrated in one information system. However, supply and demand data may not be compatible, as they are frequently collected independently of each other. Normally only some *performance* data will be attached to the network data (supply) from the national and mode-specific data providers. Such data will most probably be registered at the location and therefore by nature linked to the network, whereas demand data on the other hand, are commonly registered and modelled separately and are thus not compatible with performance related to network data.

Performance data may typically be the number of vehicles (adt or pcu) and axle passing on roads, movements and landings on airports, docking in seaports, the number of trains passing railway stations etc. Demand data are normally O-D-flows of goods and of passengers, transport chains and socio-economic data. So, supply data give information on transport capacity, services and performance, whereas data on demand give information on mobility patterns. However, separately these quantities say little about the relationship between supply and demand.

At the conceptual level, compatibility between supply and demand data would make it possible not only to monitor to what degree the infrastructure capacity is utilised, but also identify and monitor the forces behind the mobility patterns from the demand side. One example is demand for transport, which probably will continue to increase in the future and the challenge to reduce the negative effects of this growth. In order to change unwanted development one needs to know the forces behind the present choices. Overall growth rates or mode share figures are indicators for monitoring the demand development, but for decision support demand figures should be much more specific and linked to the network. Making supply and demand data compatible is a step towards understanding the reasons for the present situation, and thus enables us to improve capacity utilisation of the European transport networks.

Many transport models rest upon a representation of flows assigned to a network. Compatibility is needed when performing scenario exercises in a model, for instance estimating the modal split before and after single, or mixtures of, policy measures. Such measures could either be induced changes on the demand side (e.g. taxes on fuel) or on the supply-side (e.g. building infrastructure) of transport, and the output would be resulting changes to capacity exploitation, emissions, congestion, noise etc. One example of such a model is the STEMM model. Due to flaws in existing data, it was necessary to perform a new survey on freight transport flows between the U.K. and continental Europe, and surveys of operators in the Trans-Alpine region, in the Nordic countries, Ireland and the U.K. An assignment exercise was then called for, in order to create some kind of compatibility between the flows and the selected corridors. Without compatibility between supply and demand data, such calculations are of questionable use.

The optimal situation, of course, would be that available data were compatible and that high quality assignment could be performed through a user-friendly access to harmonised data. The model-based approach described above, is not sufficient to secure such compatibility in the long run, since different models demand different data and create different results. The guide towards the future database methodology should instead base itself on some more stable characteristics of the relationship between society and transport, like the novel statistical concepts that were established in the INFOSTAT project. The transport chain concept, for instance, was considered fundamental in order to catch the complete sequence of transport modes from origin to destination, explaining the forces behind the freight pattern much better than the purely single-modal approach. The indicators from INFOSTAT are the departure points for MESUDEMO. Thus, these long-term considerations are included in our recommendations. This is the reason why WP2 describes the road towards transport chains and why WP3 deals with methods for describing O-D flows of passenger transport. These data will have a geographical anchorage, so that they are compatible with network and socio-economic data.

So, there should be compatibility, but the detailed linking of data is resource demanding. It is not realistic to aim at having ready-to-use transport chains on say NUTS 3 level for all Europe, truly attached to every link used on the way between origin and final destination. The challenge is to make an informed choice on what detailed statistical level to operate in the general database, based on the experiences in MESUDEMO.

5.2 Linking supply to demand

Narrowing the challenge of compatibility to the task of attaching already *existing* information on flows to the networks, that is linking demand to supply technically, MESUDEMO has analysed two possible solutions. The first one is to make use of the network as an integrator of the data, and to let a GIS be the backbone of the database. This is the recommendation developed through the Dutch pilots in MESUDEMO. The other solution exploits the fact that information in ETIS is necessarily georeferenced. But, there are two main axes of organisation: The regioning scheme (e.g. NUTS 2 or NUTS 3 or any regioning scheme of variable granularity according to the problem at hand) and the network of the transport links. A unified model has been developed for ETIS, for handling all these different kinds of data, and as a general rule, all data are recommended organised in a contemporary, advanced DBMS (Database Management System), instead of a GIS. We will here only briefly present the two approaches, along the horizontal dimension in figure 5-1.

5.2.1 Linking with the network as an integrator

The principles presented here for linking flows and infrastructure data are developed from the Dutch case of the National Infrastructure Network of the Dutch Ministry of Transport. Similar ideas have been proposed elsewhere, notably in Assembling and Bridges. In these projects the GTF (General Transportation-data Format developed by Mkmetric) works along the same principles.

The recommendation for technically linking transport demand and supply data is based upon a logical reasoning: Infrastructure networks contain a spatial component due to the fact that both traffic and transportation are closely related to geographic locations and the relations between those locations. This matter of fact is exploited for the task of integrating data.

Within the network it is recommended that there are three basic inter-connected elements; node, link and region (zone). The attribute data (short for information linked to the network) should be linked to one of those three elements. In addition to the normal harmonisation and unification process of the data themselves then, also a harmonisation and unification process of the geographical definitions should be done. Only the attributes needed for identification and connection of data will be within the “ETIS”- network. All the other attribute data are stored in other databases.

There can be either a *hard* or a *soft* connection between attribute data and the network. With a hard connection the data will be attached to the link, the point or the region with a database pointer. This is the most preferable method, but it will not always be possible. A soft connection is a connection with so-called ‘spatial keys’ that should be maintained both in the network and in the other databases. The problem with these spatial keys is the maintenance, in case there is a risk of inconsistency.
The GIS orientation of transport statistics does have significant impacts on the organisation of the statistical information, in that data should be spatially related. The backbone of a GIS-based system is the representation (or better maps) of the various infrastructures to which data can be assigned. A crucial component in ETIS then, will be the maps of these infrastructure

networks. There are substantial advantages in having only one common referential structure to which data can be related. In this way many types of data sets can be related to each other through the spatial linkages and thus enabling complex relational analyses.

5.2.2 Linking in the visible plane

Usually, a GIS is used for the organisation of the infrastructure-related data. The transport networks, specifically the nodes and the links, are used as the elements, to which all other information is linked. Each mode is usually on a different plane, called layer: One for the road, one for the rail, one for the inland waterways, etc. In such a case, the inter-linking of the different planes is indirect: through common nodes, which appear in more layers, or through their geographical superposition. The attributes of the nodes and the links are used for storing and handling network-related information. In most of the existing systems a single Level-of-Detail (LoD) is used. Moreover, the connectivity of the network is fixed. Significant changes of the LoD or a different connectivity of a network dictate the construction of a new layer in a more or less manual way: by copying some of the information and modifying the rest. For assigning partial attributes to the network links (i.e. assigning attributes to a part of a link), a method called dynamic segmentation is used.

Apparent disadvantages of the conventional methodology are the following:

- Inability to handle any polymorphy of the data.
- Lack of flexibility in the cases of necessary, partial changes (as it will be explained in the next section).
- Reduced ability to support intermodality through a multimodal network.
- Reduced ability to inter-link the data, since complex heuristic procedures are needed for this purpose.
- Lack of the ability to dynamically change the connectivity of the network, so inability to face new challenges (e.g., new units of freight transport, such as small containers, which could be transported by air as well) without the construction of a new GIS-based system.
- Usually, incompatibility with other systems of this type (mainly with already existing national or international systems, which, of course, cannot change at will).

The EUROSTAT GISCO database is organised in exactly this manner, with many layers for different information. This organisation is clearly dictated by the available capabilities of the available GIS. It has led to generalisation of this line of thought to the internal organisation of most directly or indirectly geo-referenced transport data.

A new mentality of work is needed, along with a more efficient organisation of the information system. So, instead of the "planar" organisation of the data sub-sets, the concepts of the ***visible plane*** and the ***visibility level*** of the data are introduced. They are described in detail in the next chapter.

5.3 Compatibility between concepts and data

The problems we faced in conceptually and technically harmonising existing data, partly stems from the fact that the process of ensuring compatibility between concepts and data are not made explicit, and are not adequately communicated. This results in incompatible systems with differently structured data across Europe. The Commision then, is confined to ensure compatibility between *the results* of different national conceptualisations of the transport system, influenced, among other things, by the history, geography and development of the countries that has lead to different priorities. Ideally, policy makers in Europe should instead be presented with statistics that are produced in a European transport environment in which conceptualisation and model- and theory-building are made explicit, and where all nations participate in a dialogue about the way to understand the forces at work.

The empirical, horizontal, approach limits the challenge of ensuring compatibility to harmonising and organising existing data into a system, providing European policy-makers with relevant decision support information. The diverse structure of data and different technical solutions leave us with many kinds of problems, of which this report documents some from the pilot studies in the Alps and in the Netherlands. Also the MYSTIC project provides us with information on what are the most pressing harmonisation issues when trying to construct "European" data from existing statistical sources.

When implementing the long-term policy information support system there is a need, along the vertical dimension in the figure, to aim at understanding mobility and the forces at work, and to recommend indicators that reflect reality in a more informed way than existing models and present national statistical systems can provide. The shipper-surveys performed in the MYSTIC project had this additional aim of trying to grasp the forces behind transport, rather than only monitoring it. Ensuring compatibility in this perspective presents us with the question of how we conceptualise and build theories for transport development, and how we define relevant transport policy issues.

The method in MESUDEMO, then, has been to start from existing concepts, in the form they manifest themselves through indicators in existing data sets (WP2+3, WP6+7). The indicators have been scrutinised, re-conceptualised and defined in ways, which can give them as clear and unambiguous empirical content as possible. The method is often called the *two-route strategy between the abstract and the concrete*[2]: Starting from an empirically observable phenomenon (such as private car traffic increase), one tries to identify the causal powers at work (such as new, flexible work routines, new lifestyle and consumption values), and to conceptualise the phenomenon according to this instead of according to appearances (such as correlation between number of private cars and private disposable income). This is a process of abstraction, whereby objects (the actors in the social system) are defined according to certain attributes that are essential to us (individual responses to the social shaping of values) explaining the phenomenon in question (transport development). However, the objects do not appear empirically in this abstract form. Therefore it is necessary to transform concepts into indicators and variables that can be given content in the form of "data".

[2] Sayer: "Method in social science", 1992

Figure 5-2: ***Two-route strategy between the abstract and the concrete***

As visualised in the figure above, when not stated explicitly, we may all possess different opinions concerning the necessary relationships between the constituent elements of the transport system (one may believe it to consist of “Abstract concepts 1” or “Abstract concepts 2” or any other set). This may, however, lead to totally different sets of new data variables. Although all we want to do is to move from existing data variables to new, improved ones, where we end up (in “New data variables 1” or “ New data variables 2” or any other), hinges crucially on how we conceptualise the constitutive elements of the system.

Conceptualising European-scale transport development in a way that will both reveal its driving forces as well as describe its effects, is a formidable task, and there are good reasons to take first things first to get a grip on the extent of the phenomenon. In order to achieve some idea of European transport performance, harmonised data sets based on the most central concepts are badly needed, at least to get an overview of what we are dealing with. However, the ambitions of our societies go well beyond this: we want to influence the development. This, however, cannot be done solely on the basis of knowledge of appearances, but must, in order to avoid ”blaming the patient for his disease”, include knowledge about cause-effect relationships.

5.4 Recommendation

MESUDEMO recommends that compatibility between supply and demand data should be ensured in the general databases. There are two main roads towards such compatibility.

One is to design an information system in which already existing, differently structured data can be organised so that there are links between supply and demand data. This chapter briefly presents two possible models for this main road, one of them using GIS as the backbone of the system, and the other one linking data either through the regioning scheme or through the networks.

On the other main road our recommendation is to continue to encourage an explicit conceptualising and theory-building process in the European intellectual environment, so that not only do we understand mobility better, but we also understand it in a similar manner across Europe. The projects in the 4th Framework Programme, of course, have all worked on this European level, with partners from many countries trying to grasp European transport development together. MESUDEMO has participated and contributed in this dialogue, through two workshops, in-depth interviews, questionnaire survey, many meetings with external participants and general dissemination of the work.

Conceptualisation should become explicit through the implementation and use of ETIS. The proposed architecture of ETIS allows the system to function as a mediator in a dialogue, none of the modules in the architecture are thought to be automatic and static, but should rather be considered as arenas for the converging and learning process towards understanding mobility.

6 OVERALL AND INTERNAL ARCHITECTURE FOR ETIS AND THE DEMONSTRATORS

6.1 Objectives of the work

The main objectives of the work on the data architecture in MESUDEMO (WP 4 and WP8) is the following (from the inception report):

> Development of *a set of data models* and the corresponding *data architecture* for the creation of European databases on transport flows and transport infrastructure.

Moreover, the proposed *data and overall architecture* should be suitable for a future European Transport Information System for Policy Support (ETIS).

6.2 The challenge of building an ETIS – The results of the work on the ETIS architecture

Relatively early in the work of the project, it became clear that there are considerable difficulties when trying to propose an overall and inernal architecture for ETIS. Some of the most important of them will be explained in this chapter, so that the reader fully understands the differences of the ETIS case to the conventional transport databases or Decision Support Systems (DSS).

Top-down design: Not possible alone

In creating and using a national system, a *top-down approach* can be used:

> *Policy issues → Formulation of the questions, to which the information system (IS) has to provide answers → Decisions on what kind of data and algorithms have to be included in the system → Collection of the necessary data and construction of the system → Design and realisation of the information system; answering the policy options*

In an ETIS, however, the situation is different. In many cases, data are already available through agreements with the European countries. These data, however, are not necessarily interoperable, or in the same form or format. In other cases, fundamental data are either missing, or very difficult or not possible to merge. On the other hand, it also became clear in the project that the national bodies are sensibly reluctant to co-finance ETIS initiatives for acquiring new data and proceeding to further developments of this kind, because of the subsidiarity principle.

The aforementioned facts suggest the use of a different methodology of work for the creation of ETIS: the top-down design approach has to be combined with a bottom-up one, in order to take into account the *existing* data:

> *Top-down part of the design of ETIS:*

Policy issues → Formulation of the questions, to which the IS has to provide answers → Initial decisions on the kind of desired data and algorithms that have to be included in the system

Bottom-up part of the design of ETIS:
Determination and harmonisation of the available data or the data that can be collected → Given these data, determination of the possible algorithms and processing methods that can be successfully used → Identification of the questions, to which the resulting information system can answer → Design and realisation of the information system

Therefore, the actual procedure of designing and realising an ETIS, comprises two stages: *a problem-driven, top-down initial stage*, which must be complemented by a *data-driven, bottom-up one*, which will determine which parts of the top-down one can actually be implemented. It must be stressed that it might be necessary that the aforementioned procedure be repeated iteratively, even more than once, in order to reach a plausible data arrangement for working on the corresponding problems.

Moreover, in many cases the validation, harmonisation and merging of data originating from different national and international sources poses new and non-trivial problems. In such a case, *it may be necessary to use a pilot action for understanding and guiding any further work*. A very practical example of this has been the Transalpine Pilot Study, which is presented in chapter 4 of this report.

The aforementioned peculiarities definitely influence any attempt to propose an overall and internal architecture for ETIS.

The multiplicity of the transport databases

As far as the transport databases are concerned, there is a considerable multiplicity of them in Europe; *however, there has been no significant effort for a basic harmonisation of their content or for their co-operation*. The description of these databases is beyond the scope of this document. The reader is encouraged to look at the Deliverables D1, D2, D3, D5, D6, D7 and D10 of MESUDEMO (see chapter on references) for descriptions of such databases. At this point, some examples of such databases will be referenced, in order to indicate the different kinds of them:

<u>Types of existing transport databases:</u>

- Local (regional) databases, such as the following:
  - Alpine regions
  - Adriatic corridor
  - Pyrenees crossing

  and others.
- National databases, mainly created and maintained by:
  - Statistical offices
  - Infrastructure administration offices
  - Ministries of transport
  - Local administrations
  - Private companies
- International (supra-national) databases, created and maintained by:
  - EUROSTAT

- International Organisations such as the UIC, the International Road Federation, ECMT, IATA, etc.
- International efforts and studies, such as the RADEF database, the CAFT database, etc

As an example, the databases used for the Transalpine Pilot Study are referenced (please see section 4.2 of this report).

As already mentioned, the sheer number of these databases, and the fact that everyone of them has its own definitions, internal data architecture etc., emphasises the difficulty in designing an architecture for ETIS, which has to co-operate and use data from them.

The situation when the work on the overall and internal ETIS architecture started

At the starting point of the MESUDEMO work, there was a considerable amount of work going on for the development of systems and tools for processing transport information at the national, or even the European level. However, the transport experts had the definite feeling that ETIS had not been achieved yet. Ever worse, it seemed that *there was a general lack of understanding of the* ***aim****,* ***coverage*** *and* ***form*** *of ETIS*, due to the difficulty to grasp the organisational, institutional and technical concepts of such an extended and complex system.

So, important questions were still open, such as: Exactly which problems are ETIS going to tackle? What is the true meaning of the word "European" in ETIS? Where is ETIS going to reside? What kind of software and hardware platforms and tools are needed for the implementation of ETIS?

Some of the most characteristic organisational and technical issues, concerning ETIS and the databases for networks, which were still open at the starting point of MESUDEMO, are presented here:

- What *kinds of information* have to be included in an ETIS? *At what level of detail* do they have to be? How can different levels of detail co-exist in a single system?
- *Which GIS must the Commission use*? (EUROSTAT uses ArcInfo in GISCO, a number of organisations (e.g. UIC, the Italian MoT, etc.) and companies use MapInfo, the Commission has a proprietary GIS, GEOSYSTRANS has developed new solutions, etc.) – *How can GIS compatibility be achieved*?
- What kind of connections of ETIS to national and international databases are needed? How can they be implemented? How can the data acquisition paths be properly organised and operated on a regular basis?
- How can we store *knowledge* (know-how) in the system and relate it to the other stored information?
- How can *data confidentiality* be handled? What about *data privatisation*?
- How can the data quality *certification* be handled?
- *Where is ETIS going to reside?* Is it going to be *a single system*, at a certain location in Europe, *or not*?
- What is the *overall architecture* of an ETIS-type system?
- In which way have the ETIS data to be internally organised, so that the practical system is *open and scalable*, both as far as the data and the processing methods are concerned?

- How can *data re-usability* be achieved (without, of course, building another system from scratch, each time we need to modify only a part of the data, but, instead, extending the same system and using it for both applications)?
- Which are the *software and hardware tools* necessary for the ETIS development?
- *Who is going to pay* for the development of ETIS? Who is going to pay for the maintenance of ETIS and the data acquisition?
- Is there a practical, realisable path towards ETIS?

In fact, at the starting point of the MESUDEMO work, it was extremely difficult to answer even the following fundamental questions:

> Suppose that the funds were available for starting the implementation of ETIS. *Which are the first steps for creating ETIS? Is there any preparatory work that has to done first?*

At the same time, on the other hand, a number of crucial changes were happening during the life-cycle of the project, in the following areas:

- *In the development of Transport Information Systems (TIS)*: In many cases, there had been a definite turn towards *privatisation*, while the simultaneous development of similar and antagonistic systems had already triggered a phase of competition (instead of the previous monopolies). At the same time, a lot of new TIS were emerging or under construction in the Member States (national or private), in international organisations, etc. It is apparent that these changes at the *data sources* of ETIS do reflect in the architecture and design of ETIS.
- *In the area of information Technology and Telecommunications:* The commercially available Database Management Systems (DBMS) and the Geographical Information Systems (GIS) were decisively evolving and considerably increasing the possibilities for the users (e.g., by enriching their internal data architectures, providing the means for creating *open and scalable systems*, allowing the creation of *web-enabled and portable* systems, allow the *federation of databases*, etc.). The proposition for an architecture for ETIS must definitely take into account this very fast evolution.
- *In the very transport policy implementation of the EU:* There had been a decisive turn from the centralised approach to the assignment of regional responsibilities, as well as a lessening of the number of new directives and an evolution of the policy issues of interest from administrative and transport-centric to citizen-centric. This actually means that even the questions, to which an ETIS has to provide answers, may change very rapidly.

All these changes had direct and decisive effects on the design of an ETIS. The MESUDEMO had to provide architectures, which must be able to cope with this constantly changing and evolving environment. The only way to do so, was

> *to look at the broader picture,* and instead of defining and designing a *static, closed ETIS*, examine the possibility of *designing and producing ETIS* as a *set of methodologies and open and adaptive computational tools* for investigating the impacts of policy and transport options or measures, *able to cope with the constant, decisive changes.*

This concept has actually been a guideline for the whole work on the ETIS architecture.

Further challenges faced in MESUDEMO

In the evolution of the work on the overall and internal architecture of ETIS in the MESUDEMO project, it has been realised that a number of peculiarities or specific needs in the ETIS case dictated specific design and implementation methods and/or the introduction of innovative solutions. Typical issues of this kind include:

- the *expected polymorphy and diversity of the data* in an ETIS
- the concept of the *ETIS network* (i.e. a network of interconnected, co-operating systems, distributed over Europe, as opposed to a single, huge system situated at a certain European location),
- the need for seamlessly including *proprietary data* in the data core of ETIS,
- the ability of ETIS to co-operate with any GIS at any time (*GIS independence*), and others.

The MESUDEMO consortium had to face the challenge and provide feasible and practical answers and novel solutions to all these questions and problems. The aforementioned peculiarities and specific needs of ETIS, as well as the practical answers and novel solutions provided by MESUDEMO, will be described later on in this chapter.

6.3 The results of the work on the ETIS architecture

In the following, **the main results of the work on the ETIS overall and internal architecture** that have been achieved in MESUDEMO are outlined:

- **A better understanding of the organisational and technical nature of ETIS has been achieved** (aim and form, technical needs and solutions, installation and maintenance, possible problems still remaining, etc.)
- **A concrete need for preparatory work at the European level has been identified**, and corresponding recommendations and suggestions have been made to the Commission.
- **A general, overall architecture of an ETIS-type system has been developed**, and the main modules of it have been identified and further analysed.
- **A novel, unified data architecture and the corresponding data model for the infrastructure networks and the flows have been developed**. This architecture:
  - *Starts from the novel concept of issue-relevance* (i.e. the relevance of the associated information to the transport or policy issues at hand), *in order to cope with the polymorphy and diversity of the data.*
  - For the first time, takes into account critical parameters like the *Level-of-Detail*, the *possible locality of the problem at hand* (e.g. the local problems of specific corridors), etc.
  - Can support *widely different needs*.
  - Can use *any GIS at any time* (i.e. it is GIS-independent).
  - *Permits the seamless integration of external, third-party data sets and processing methods* (including models), without violating the ownership and confidentiality of them.
  - Follows the current State-of-Art for creating *open and scalable systems*.
  - Can *anticipate future needs* for different organisation of data sets of the system.

- A *realisation methodology* for ETIS, *which reduces the installation and maintenance costs and permits the open (but controlled) access to its data and processing methods,* has been developed.
- *A step-wise, practical path to a future ETIS has been envisaged and described.*

In the following, these results are presented in brief.

6.4 The organisation and technical nature of ETIS

In the beginning of the MESUDEMO project, ETIS was being thought as a large system, residing at a certain place in Europe (such as EUROSTAT or DGTREN) in which the proper information from all Europe would be imported and used for the needs of policy support. One of the first ideas was that ETIS should be the “common denominator” behind the national systems or other international systems.

However, during the evolution of MESUDEMO, it has been realised that this “centralised system” was not really the best option for all Europe. Instead:
- Specific, *localised ETIS sub-systems* may be built for dealing with problems and political issues of multi-national character, which, though, are particular for certain areas. Such an example is the ATIS (Alpine Transport Information System for policy support), studied in the Transalpine Pilot Study in MESUDEMO. This statement is justified by the fact that in many important cases, both the transport problems and the policy issues of importance may be different (or may be expressed in a different way) for different geographic areas. Therefore, the subsidiarity principle dictates the creation of localised ETIS sub-systems.
- Actually, *an ETIS system may even appear at the national level*, if a national TIS or DSS contains an adequate amount of multi-national data, in order to allow the study of transport situations and problems in bordering multi-national areas (e.g. for study of the Turin-Lyon corridor).
- Moreover, a different ETIS system would be *one that would be used for transport planning at the European level and is covering all Europe*, or at least a large part of it; or a system that is used for studying of European-wide situations or problems.

Therefore, *a possible and, actually, probable form of ETIS would be that of a network of interconnected, co-operating systems*, situated in various places around Europe and forming what has been called in MESUDEMO ***the ETIS network***. It is not necessary to integrate these systems; the interconnection of them can be as tight or as loose as necessary or meaningful. These systems, however, have to be designed beforehand in such a way, as to co-operate in a seamless manner, if and when such a need arises. Databases of this kind are called *federated databases*.

Moreover, it has also been realised that a *special architecture* is necessary in order to cope with the following demands:

- The existence of *very large volumes* of data.
- The achievement of *the maximum re-use of the data*, without building new systems from scratch each time a sub-set of the data has to be used.
- The achievement of *data interoperability* to the maximum possible extent.
- The *easy, efficient and less costly interconnection* (federation) of the ETIS sub-systems.
- The *easy and less costly maintenance* of the system.

The principles of such an architecture have been developed in MESUDEMO and are described later on.

6.5 A general overall architecture for an ETIS

A general skeleton of a Transport Information System for policy support can be seen in figure 6-1, where the main modules and their basic functions are depicted.

Figure 6-1: A general architecture of an ETIS-type system, showing the various sub-systems at a conceptual level

The general conceptual modules of the system are depicted in figure 6-1 and explained in detail in the following. One must clearly distinguish between the conceptual modules of the system, as they are presented here, and the possible actual components (programmes or blocks of them), which will be used for the implementation of these modules.

In order to better understand the architectural and technical issues, which are related to each module, it is very important to first identify the *functionality* of the module, and then examine *the best technical ways to implement the related, actual system components*. Then, and only then, it will be possible to propose an internal architecture for the components used for the implementation of the system.

The supportive module for the data acquisition, harmonisation and combination procedure

The main discriminating difference of an ETIS from any similar national TIS or DSS is the incorporation of data, which come from different (mainly) international sources. In order to use these data, a procedure, which may probably involve harmonisation, validation, aggregation, merging, etc., must be followed. In most cases, the intervention of experts, who will make critical decisions or choices, or who will, if necessary, suitably modify this procedure, is still mandatory. On top of this, in many important cases, the possible *lack of data* will make the work of human experts even more necessary.

In cases where the data acquisition procedure will be used a number of times for regular updates of the data core of the system, *selected re-usable parts of the procedure can be automated*. In that sense, a module can be created for the support of this re-usable part of the procedure having as a background the development of a “learning” system. This is the *supportive module for the data acquisition and combination procedure of the presented skeleton* of ETIS. Moreover, this module can actually contain a kind of workspace, in which the experts and the professionals, who are performing the data combination procedure, can perform necessary tests, transformations, merging etc. on the data, in order to reach a consistent data set, usable in the data core of the system.

The module for preservation of knowledge and data: The concepts of the transport data warehouse and the knowledge base

The original data sets, used to create the data core of the system, should be kept available, so that corrections or adjustments can be made to the data core of the system at will, and new sets of combined data may be added. The system should provide such a possibility, therefore the introduction of a data warehouse in the system modules in the general ETIS skeleton. The information does not necessarily have to be a part of the highly organised data core of the ETIS-type system. It may be in several formats, loosely connected to other information, but still it may prove quite valuable for the expert users. Therefore, an ETIS-type system should include both a *data warehouse*, in which original data sets can be efficiently kept, and a way to efficiently maintain unstructured information (texts, multimedia hypertexts, etc.) and knowledge (know-how) through the use of a *knowledge base*.

An additional, but very important, reason for the inclusion of a knowledge base (and a data warehouse, for that matter) in the overall ETIS architecture is the following: The process of creating such a system is very complex and involves many necessary decisions, taken at various stages of the feasibility study, the design and the implementation of the system. Many of these decisions regard the selection, combination and use of the data themselves.

As far as the data warehouse is concerned, the critical issues concerning its operation are the following:

- The ability to store (and possibly handle) a large number of data sets in their native formats.
- The provision of a kind of special links, called issue-relevant links and indices, (and described in the Deliverables D4 & D8), which relate the stored data to the problems at hand. These links guarantee that the data sets of the warehouse can be referenced from the other modules of an ETIS at will.
- The ability to seamlessly access on demand data stored in remote databases; in that sense, it is a *virtual warehouse*. This operation should be transparent to the user to the maximum possible extent.
- The ability to provide the expert users with suitably designed data search and correlation methods and utilities, so that they can efficiently look up for the existence of specific data.

As far as the knowledge base is concerned, the stored (or accessible) information could be in the form of textual descriptions and explanations, methodologies, algorithms (e.g. computational methods, models, etc.), examples, etc. Apparently, the information of the knowledge base has to be thematically connected to the meta-information, which will accompany any datasets, so that it can be directly accessible by the users of the data. Actually, the information of the knowledge base can be considered, in a sense, as a more extended and thorough meta-information.

The highly organised data core

Any decision support system, in order to compute the indicators and/or perform the simulations or predictions and "what-if" tests, which are needed for policy support, must contain a data core. This data core, constructed according to the aims of the system, includes the information necessary to feed the set of inference engines (processing methods, models, etc.) of the DSS. In ETIS, the data core usually includes infrastructure information (mainly the transport networks), demand data (flows), as well as socio-economic, behavioural, environmental and other region-related or network-related information. In contrast with the transport data warehouse, where the data can be loosely inter-linked, there is a strong need for highly organising and inter-relating the data of the data core. Moreover, it is possible that the data core may contain organised information, which is not necessarily fed to any inference engine pre-included in the system, but may directly enable experts form an opinion on the evaluation of critical policy options.

In MESUDEMO, several architectural schemes have been examined, as well as the international and European state-of-art. Moreover, many experts have provided their opinion. Having this as input, the MESUDEMO team has developed a novel and advanced architecture for the Data Core, with many advantages when compared to the existing systems.

The module for the extraction of information from the data

This module is actually what differentiates a Transport Information System (TIS) from a Decision Support System (DSS). It comprises a number of inference engines, which produce the indicators or perform simulations or predictions, needed for the support of policy-making, based on the data of the data core. These inference engines can range from simple data extractors and viewers to very complex model chains. They usually include methods for processing the data of the data core, including transport (or other) models; they also have to supply utilities for scenarios formulation and hypotheses formulation and testing, as well as additional workspaces for the temporary storage of the organised outcome of extended projections or predictions.

It is important that this module also includes facilities for *monitoring the area of coverage on policy and transport issues*, as well as any related directives, regulation, studies, constraints (time, speed), etc. Adequate links to the knowledge base and the data warehouse must be provided for this purpose.

The transport-oriented and policy-oriented user interface and the user guidance module

The quality of the user interface is important for the easy, efficient and correct use of the system. Moreover, it is a given, strict requirement that the ETIS interfaces share at least a common "look and feel", if not a common way of handling similar situations, so that as limited a training as possible is required for the users, when the latter use the various interfaces or sub-systems. However, *not all interface needs can be anticipated for every ETIS sub-system*, since they may considerably vary according to the problems, which the sub-system may handle. Therefore, a common *interface design methodology* for ETIS has to be developed. This methodology should end up in a set of design methods, as well as templates and tools for building the user interfaces needed for any ETIS system. This methodology must also cover to the maximum possible extent the necessary variability which is introduced by the use of different components, such as models, GIS, etc. These components may be different, but the interfacing methodology must be the same. *The development of this methodology should be one of the first jobs towards ETIS, and should be driven by a set of pilots*. Using the resulting set of templates and tools, a first set of user interfaces for the most common cases of each ETIS sub-system has to be developed.

The user interface modules may include: *Data or result viewers* (for tables, data forms, text or hypertext, geo-data, graphical forms like graphs, pie charts, etc. and other visualisations), *interactive modules for communicating with the users and, most importantly, guiding them*, *modules for producing reports, etc.*

It is obvious that the development of the interface design methodology and the realisation of the first set of user interface and guidance modules can only be done, if the specifications of an ETIS has been first determined and the design of ETIS is under way.

Extra care must be taken, so that open and widely accepted tools and methods are used, in order to allow the easy expansion of the system; *it is suggested that a three-tier logical architecture*, which implements separately the database tier (usually called data servers), the application tier (i.e. the way that the data are retrieved and processed, according to the needs of the individual problem at hand) and the user interface tier (i.e. which handles the way that the information is presented to the user and the interaction of the user and the system) *is used for ETIS*. In this way, *openness and scalability is maintained and the re-usability of ETIS is facilitated.*

The set of mediation modules (portals to and from other systems)

Whenever there is a need for repeatable, systematic communication of ETIS to other organised systems (i.e. databases, like the EUROSTAT databases, databases kept by the Member States or international organisations), special modules of the system will handle the data transfer. These modules will act as a portal to (and from) other systems and automate to a maximum possible degree these operations, in order to avoid duplication of effort, each time such a data transfer happens. The modules will make certain that a properly structured transfer will occur, i.e., concerning the transfer of data from external systems to ETIS:

a) That the right data are extracted from the external system, that they are transferred to ETIS and properly validated.
b) If necessary, that the transferred data are properly transformed (to quantities usable by ETIS).
c) That the transferred data are accompanied by the proper meta-information and documentation.
d) That these data will be fed to the right ETIS modules (e.g. the supportive module for the data combination procedure, the data core, etc.), or stored in the right place (data warehouse, highly organised data core, etc.)

On the other hand, concerning the transfer of data *from ETIS to external systems*, functions similar to those of (a) to (c) of the above operations will be carried out.

It must be highlighted that the *mediation* between systems, when data have to be transformed, computed or merged, so that they can be properly used by other systems, is a way to achieve *data interoperability*. Therefore, mediation can be thought as a, temporary at least, *alternative to standardisation*.

Two additional functions of this module are important and necessary: an efficient co-operation with a *meta-information server of ETIS for handling virtual data* (i.e. data stored in other systems but functionally integrated in ETIS), as well as an *efficient handling of any data confidentiality or ownership issues*, again in co-operation with the proper ETIS logical modules.

The issue-relevance determination module

One of the major problems in the systems of the ETIS-type is the *multiplicity of the possible forms* (*polymorphy*, please see the glossary) *and formats* and *the diversity* of the involved information. It will be a common phenomenon that different internal representations of the same entities (such as the transport network links and nodes, etc.) may be needed in order to tackle different problems (examples for this are presented in the deliverables D4 and D8). *Drastically restricting the possible internal representations of the data,* in order to reduce the effects of this problem, *is clearly unacceptable*; *it might result in an inflexible, or even inadequate, data core.* Moreover, *it will severely limit the re-usability of the data sets contained in the data cor*e. In MESUDEMO, a novel methodology for tackling this problem through a *conceptual sorting* of these polymorph and diverse data has been developed. This methodology allows such data to co-exist in the same system without any confusion. It relies on the use of an *issue-relevance determination module*, which bases its operation of the relevance of a sub-set of data to the problem at hand (i.e. the transport or policy issue), in order to organise internally the data and select the proper sub-sets needed for tackling each problem.

Moreover, this module properly relates the data to selected meta-information and to the knowledge, which is stored in the knowledge base of ETIS. Although this functional module of ETIS is not at the same level as the modules presented in the conceptual skeleton of ETIS, its significance is crucial for the proper organisation of ETIS, and it is underlying the whole proposed ETIS internal architecture.

6.6 A novel, unified, internal data architecture and the corresponding data model for the infrastructure networks and the flows

Due to the highly technical nature and the extent of the content of this section, the results will necessarily be presented at an abstracted level.

The different kinds of data in ETIS – ways to organise them

The ETIS data can be split in two main categories:

- (*Structured) transport data*, with the usual meaning that the transport experts assign to this term, and
- *Unstructured data*, usually textual or multimedia data, including documentation, EU or national directives, knowledge representation, etc.

The structured transport data can then be further split in:

- *Static (or semi-static) data*, i.e. data that regularly do not change or change very slowly, as far as the system is concerned. Usually, most of the infrastructure-related data as well as socio-economic data fall in this category, since they change (i.e. they are updated) very slowly.
- *Dynamic data*, i.e. data that change very fast and need special methods for handling. E.g., such data are the traffic data on a road network, for which updates may be necessary very often.
- *Time series*, in which the information lies in the evolution of certain indicators or quantities over time.

The static or semi-static data are further split in two main categories, according to the way they are organised:

- *Infrastructure-related (mainly network-related) data*. All data that are usually associated with the infrastructure networks (e.g. transport links and nodes, their attributes, partial attributes and their geometry, services, routes, etc.) fall in this category. Even counts of flows on the links fall in this category. The kind of information associated to the networks may change significantly with the Level-of-Detail of each individual network.
- *Region-related data*. These are mainly demand, socio-economic or behavioural data, associated to regions. Again, the kind of information associated to the regions may change significantly with the Level-of-Detail of each individual regioning scheme. For the regions, as well known, there is already a European standard, the NUTS zoning.

Usually, a Geographical Information System (GIS) is used for the organisation of the infrastructure-related data. The transport networks, specifically the nodes and the links, are used as the elements, to which all other information is linked. Each mode is usually on a different plane, called layer: One for the road, one for the rail, one for the inland waterways, etc. The work in MESUDEMO, however, has shown that such a conventional organisation can prove both inefficient and restricting for ETIS.

As far as the *flow data* are concerned, more than one organisational scheme are usually used, depending on the form of the information, which usually is in the form of:

- Trip data, at various levels of aggregation. Relational databases are used for this purpose, usually not geo-referenced. Geo-referencing is implicit, through the origin and the destination of the trip.
- Origin-Destination (O-D) matrices, in various forms and levels of detail. The corresponding databases can have different forms, although they remain basically relational databases with indirect geo-referencing.
- Counts on the links of a network: actually, although these data concern flows, they are observed data on the links of the networks. They are usually organised in the network part of the systems.

Of course, if one interprets the "databases for flows" in a little broader sense as *"databases, which describe the transport demand"*, then other kinds of data, such as *socio-economical*, *behavioural*, etc., have to be considered as a fundamental part of such databases as well. These data are usually organised again as region-related, relational databases, and can take a multiplicity of forms.

In MESUDEMO, a unified model has been developed for ETIS, for handling all these different kinds of data, as well as the necessary *meta-information*. As a general rule, *all data have to be organised in a contemporary, advanced DBMS* (Database Management System). The need for geographical visualisations and spatial analyses are covered through the use of a *spatial data interchange module*, which takes care of the communication of the database to a GIS (such modules have already become commercially available). *Multi-scale information may or must be present in an ETIS*; a special *Level-of-Detail (LoD) stamping* of the data is needed, in order to be able to mix information with *variable granularity*, i.e. of different LoD. This mixing must happen on a kind of *visible plane*, which is the actual working space containing the information needed for a specific study, as the following figure shows:

Figure 6-2: The relationship of the visible plane to the conventional layers of a GIS

Moreover, except from the LoD stamping, the data has to be properly prepared, so that the visible plane can be effectively used. This is a preparatory work for ETIS, on which the Commission should take an initiative.

The concept and methodology of the " issue-relevant organisation"

As already mentioned in the ETIS overall architecture, the very reasonable and desirable effort to re-use in any possible way the data of an ETIS, therefore the effort to use the same system(s) for the study of more than one problem, results in an *increased diversity and polymorphy of the data of the corresponding data core(s)*. The term *polymorphy* is used to signify the phenomenon, where different internal representations of the data may be (and usually are) needed, according to problem or issue at hand. In that sense, *the form and the internal representation of the data of an ETIS (actually of any IS) are issue-relevant*, i.e. they absolutely depend on the issue (policy, transport, socio-economic issue or option, or any problem), on which an answer has to be provided.

In order to cope with this polymorphy and diversity of the data core of ETIS, a novel internal organisation is proposed for the first time:

> *Not only the data themselves, but the very issues or problems, for the tackling of which the data can be used, should be stored in the IS and used for a kind of "conceptual sorting" of the data.*

This is done through the use of a special module called *issue-relevance determination module*. The existence of this module and the use of it for conceptually sorting the data of the data core permits the co-existence of very complex, diverse and polymorph data in the core; the user may obtain from such an organised core *custom views of the datasets, highly tuned to the issues or the problems he studies*.

Therefore, with such an organisation in an ETIS sub-system or in the ETIS network, more than one specific or general networks (e.g. at different Levels-Of-Detail, or with variable granularity) and regioning schemes may co-exist and may be used as the heart of the organisation of the data of the data core(s), without sacrificing the efficiency of the system and the easiness of handling the corresponding data. The following figure shows the basic principle of this organisation:

Figure 6-3: Dependence of the network and regioning scheme information (co-existing networks and regioning schemes) on the issue-relevance module

The data model for the networks, the regions and any associated information

The internal representation of both the networks and the regions (or sub-network or sub-region) should start from an abstract, not *geo-referenced,* internal representation, called in the figure "an elemental representation", which uses a *unique, unambiguous coding scheme* for the elements of the infrastructure or the regions. Although such a scheme exists for the regions (the NUTS nomenclature), there is no equivalent for the networks. It is imperative that:

- A *harmonised data dictionary* should be created *at the European level*, in co-operation with the Member States, *for the most significant network and region elements, for all the Levels-of-Detail of interest*.
- Moreover, *a hierarchical coding scheme for clearly and unambiguously identifying each network link and node, from the EU level down to the national level, must be created.*

Again, *this is a concrete part of preparatory work for ETIS*, for which the initiative belongs to the Commission.

The elemental representations must not be based on geo-referencing, since practice has repeatedly shown that, if geo-referencing is used for identifying elements of the infrastructure, small errors in defining the position of the elements may create huge discrepancies in the representations.

Creating pan-European networks is not a trivial task; the information must come from those who produce and maintain it, namely (and mainly) the Member States. DG TREN has to use at the European level a multi-scale network. I.e., information in more scales than one (three are considered adequate) must be kept in the European networks databases and ETIS, and, each time, a specific scale must be selected according to the question at hand. All scales of the networks should be acquired in each update. Again, the aggregation from one LoD to the others should be left to those, who collect and maintain this information, namely the Member States. However, defining the needs at each selected scale is a concrete job for the Commission. I.e., in order to be able to provide the network information, the Member States have to have a clear knowledge of the kind of desired networks in the corresponding areas and the specific problems, for the solution of which the information is asked for.

Moreover, regarding any scale of the networks, or any specific sub-network: *the Commission must move towards the creation of a unique network, which includes all modes and handles them as uniformly as possible*. Special methods for creating and handling this unified networks, so that flexibility is maintained and future needs may be anticipated. Additionally, special networks for transport modelling (*modelling networks*) should be obtained and stored in ETIS as separate subnetworks.

As far as the *regions* and the associated information (transport demand) are concerned, the problems for creating *pan-European databases* are mainly conceptual and methodological rather than technical; the reader may consult the results of the Transalpine Pilot Study and MYSTIC project for further information.

In the following figure, the basic principles of the internal organisation of the ETIS Data Core are shown:

Figure 6-4: Basic principles of the internal organisation of the ETIS data core

The whole organisation is clearly issue-relevant. Special links are used for conceptually sorting of the information. The main information is organised around the "elemental representations" of one or more networks and one or more regioning schemes. Further information is linked to them, either through their coding schemes (*code-linked information*: also called *attributes*), either through another kind of "linking pins", called *keys* (*key-linked information*). The keys may be:

- *Geography-dependent*, i.e. information linked through geo-referencing of certain geographical attributes, e.g. facilities of a port residing at a certain location), or
- *Geography-independent*, in which no geographical information or geo-referencing of the element of the infrastructure, to which the key links other information, is present in the key. Usually these keys have a form, which is constituted from the code of the element plus a kind of an identifier; this identifier indicates the exact location on the infrastructure element, where the information has to be linked.

Key-linked information can be:

- Either partial attributes, i.e. attributes valid for a part of the link, such as a varying number of lanes of a highway; e.g. the number of lanes of a highway can vary as fol-lows: from the start of the link to key1, 2 lanes per side; from key1 to key2 3 lanes per side; from key2 to the end 2 lanes again.
- Or, general information; e.g. accident information. E.g.: An accident has happened at the location marked by the key [<link code>, <distance in km form the beginning of the link>].

Apparently, any information associated to a network or a region through the keys is also *is-sue-dependent*. E.g., accidents could be accurately linked to a road network at a certain Level-of-Detail, but probably not to a much more aggregated network; aggregated accident information should be used for this more abstract network.

In the proposed scheme, it is possible for the first time to actually include in the data core (and the corresponding data model) critical parameters of the data, such as the *Level-of-Detail of the data*, the possible *locality of the data*, the *availability of the data*, etc., through the use of the issue-relevance determination module and the corresponding methodology for the organisation of the data.

As already indicated, a DBMS and not a GIS has to be used for the organisation of the data. In the previous figure, the way a GIS should be used is shown: the spatial engines and viewers are included for this purpose. In the next figure, the way to connect a GIS to the ETIS data core is shown:

Figure 6-5: The way to connect a GIS to the data core

The reason, for which the use of such a module makes the architecture independent of specific GISs, is apparent: If a different GIS has to be used, simply a different SDIM (Spatial Data Interchange Module), tuned to the new GIS must be used. In this way, the ETIS data core and architecture do not change, when the GIS changes. Through the introduction of the SDIM, any commercial GIS (e.g. ArchInfo, which is used by EUROSTAT and others, MapInfo, which is used by UIC, consultancies and others, etc.) or proprietary GIS (the BRIDGES GIS, the GEOSYSTRANS GIS, etc.) or any other GIS can be seamlessly integrated to ETIS at any moment. Routes and services can be represented in a way shown in the next figure:

Figure 6-6: Handling routes and services

In fact, the routes or services should be internally represented *in a double manner*:

- First, the routes and services should be present as attributes at the links and the nodes of the networks. In this way, there is an immediate reference at the services available at any point on the network. Moreover, the operations of splitting or unifying links or regions can be checked for validity against the corresponding services and handled accordingly. Any services on *parts only* of the links of the network can be associated to the network as *partial attributes*, without introducing any unnecessary “intermediate nodes”.
- Second, a separate table of the routes, the services and their attributes should be kept, for immediate reference to the services and proper organisation of any information associated to them.

Another important issue for the data core is the *stamping of data*. Stamps are special characterisations of individual data items or clusters of data. Sometimes they can be treated as attributes of the data; however, they are mainly used to differentiate the data or the behaviour of the system accordingly. Stamping the data of the ETIS Data Core is absolutely necessary.

The need for the following stamps has been identified so far:

- A *time stamp*. This stamp permits the creation of differentiated data views, according to the time of observation of the networks or regions. Combined with a *history file*, in which the permanent alterations of the networks or the regions (or critical information associated to them) are stored, it permits advancing forth and back in time, in order to get specific future or past views of the information. Such “roll forward” or “roll backward” operations are needed, e.g., when “what if” experiments are performed and comparative studies of forecasted and present or past situations are made.
- A *Level-of-Detail stamp*. This stamp is necessary, in order to facilitate the co-existence on a visibility plane (and on a GIS) of information from different LoDs
- A *confidentiality stamp*. This stamp is necessary for enabling ETIS handle in a proper manner and resolve any confidentiality issues. The way to do so will be explained in the next section.
- A *data quality and certification stamp*. This stamp is needed for collectively characterising the quality of the data and assigning to them a possible quality certificate.

Of course, one could think of other kinds of useful stamps. Again, *the stamps are actually issue-relevant*. The way of determining the exact number and form of them is by studying the issues or problems, for which the corresponding data are needed.

Finally, in D4 and D8 the kinds *of meta-information* needed for accurately describing the data are identified, and ways to use them are hinted.

6.7 A methodology for the technical construction of ETIS

The identification and proposition of the most *open*, *overall and internal*, *architecture* for ETIS, which, at the same time, would guarantee *an easy to maintain, update and expand configuration*, has been one of the main concerns of the consortium of MESUDEMO. In this sense, a special effort has been devoted in better understanding proper realisation methods for ETIS, i.e. methods for building the actual information system. The propositions are the following:

- An *open and portable development (programming) environment* has to be used for the realisation of ETIS; such a suitable environment would be, e.g. the Java environment, due to its open character, its portability across hardware and software platforms and its low cost. However, there are also other commercial alternatives.
- Instead of a closed, proprietary system, *an open configuration has to be created, which can serve both as a model for other systems of this kind, and as a "platform" for attracting the creation of third-party "add-on" programmes and the integration of third-party data sets and processing methods (including models).*

The only restriction to the openness of the system is *the need for honouring any existing confidentiality requirements*. For any major user access to the data, ETIS actions should be along the following lines:

User request of data for a specific purpose (→ Issue-relevance determination: selection of data sets and determination of the selective and significant links) → Confidentiality Check → Check for the satisfaction of the confidentiality requirements by the users → Logical extraction of the data and granting access to them (→ Creating a physical copy of the desired dataset or database for the users)

Technical methods for implementing this procedure are proposed in D4 and D8. On the other hand, a methodology for achieving the integration in ETIS of specific data sets (or databases) and/or processing methods (including models), which have been developed by third parties, is the following: The builders of ETIS (actually of any sub-system of the ETIS network) should provide *a proper programming template* for doing so. This template would contain the programming code for the skeleton of a component (see glossary), which, if filled-in with proper programmes by the third party, would co-operate with ETIS and integrate in ETIS data sets and/or processing methods that have been created by this party. The resulting software component should include any mediation needed for the proper translation of the data between ETIS and the external systems, as well as any confidentiality resolution procedures for the access of the third party proprietary information or software.

The advantages of such a scheme are apparent: The additional cost for this procedure is split between the ETIS authorities and the third party, the third party software will necessarily be "well behaved" and the third party would be certain that the agreed requirements and procedures for accessing its data and processing methods will be fully honoured.

As far as the selection of the actual software and hardware platforms for the realisation of an ETIS system (i.e., a member of the ETIS network) is concerned, the only restriction is the use of platforms, which would guarantee an open and scalable result. Additionally, the ETIS builders should avoid the dependence on a single provider, or on products that cannot be interoperable with the most widely accepted software of today; *such dependence would limit the openness of the system and the easiness of creation of third-party software for ETIS.* Furthermore, it is advisable to choose products, for which the probability of evolution and compliance to the future State-of-Art and technological trends is high. If proprietary software parts have to be used, then they should be handled in one of the ways, which guarantee the openness and scalability of the system.

6.8 Identified preparatory work at the European level – A realistic path for the construction of an ETIS

The success of the ambitious effort of creating, or, more accurately, *deploying* ETIS (and the ETIS network) for Europe, absolutely depends on a correct understanding both of the way that such a deployment has to be done, and of the fact that a series of preparatory actions on behalf of the Commission are absolutely necessary.

These actions fall in the following categories:

- *Awareness of the social and transport problems, for which ETIS is definitely needed*:
  - There must be a very careful, comparative evaluation of current transport problems or policy issues in areas around Europe, for which a localised TIS or a DSS is a need with a high priority. The creation of a localised system for these areas has a meaning, only is it can be used in a sustained manner (and therefore, it is regularly updated).
  - Moreover, there must be a comparative evaluation of the needs for installing a Pan-European system for making possible the study of European-wide problems and support the decision making process, either directly through the use of a DSS, or indirectly, by providing information, which can help in better understanding the problems and the implications of the corresponding policy and transport options.

The reason for this essential action is *to clearly set the priorities for taking the first steps towards the deployment of ETIS*. It may be necessary to create a steering committee for deciding on these first steps.

- *Preparation* for the creation of European-wide databases for transport. A better understanding of the *data needs*, *the methods for the initial data acquisition and the regular data updates*, as well as *the form that an ETIS would take* is absolutely essential, in order to take the *correct* steps for the deployment of ETIS. The results of the Transalpine Pilot Study should be carefully investigated in all their dimensions and the resulting implications on the methodology that has to be followed, in order to go on, should be deeply understood.
- Moreover, the Commission should initiate certain preparatory actions at the EU level concerning organisational and technical matters. These actions include:
  - Promotion of an action for the *creation of a data dictionary* at the European level for a *basic common data set of transport information*. This data dictionary should be complete with an internal description of the corresponding data, not merely a glossary, so that it could be directly used for the organisation of the data core of ETIS. *This data dictionary can only be issue-relevant*, i.e., that the kind and form of its data directly and absolutely depend on the transport and policy problems that have to be faced. Therefore, *the creation of this data dictionary should be considered as a dynamic procedure*: the start should be done for a typical set of transport and policy issues, and updated and expanded in the future in a controlled and technically sound way. This update and expansion should be guided by a careful examination of the corresponding State-of-Art and the evolution of the data needs, which is perceived as the deployment and exploitation of ETIS continues.
  - *Provision of a proper nomenclature and coding schemes for essential transport information*. E.g., an action for the creation of a unique identification scheme for the infrastructure elements, which does not depend on geo-referencing, as aforementioned. *This action should be carried out in co-operation with the Member States*.

It is also understandable that in the evaluation of the overall situation, as well as in carrying out the first steps towards ETIS, other important preparatory steps may come up.

It is the opinion of the MESUDEMO consortium that the overall responsibility for getting the whole effort towards the ETIS deployment started, as well as the decision on the first steps towards this aim, should probably be taken by the primary users of the system, namely DG TREN. However, since these matters are of a highly political nature, it would be advisable to refer to the authority of the European Parliament for validating, or even directing these actions.

One of the concrete results of the architectural work in MESUDEMO has been the identification of a realistic path for the construction of an ETIS sub-system (i.e. a member of the ETIS network). The inclusion of the Transalpine Pilot Study in MESUDEMO, as well as the many contacts with national administrations and their representatives and experts, have considerably helped the MESUDEMO consortium to understand many of the details and intricacies of such a procedure.

6.9 Recommendations for the overall and internal ETIS architecture and the technical matters concerning the construction of an ETIS

In description the results from the work on the ETIS overall and internal architecture, both in this report and in the deliverables D4 and D8, a lot of suggestions and recommendations are included. Not all of them can be repeated in detail here, due to space limitations. Therefore, the basic lines of them will be abstracted and presented here, with proper references to the sections of this report or the deliverables D4 and D8, where they are better described. The recommendations are organised in thematic groups.

On the organisation and technical nature of ETIS

- The term ETIS must be associated not only with an information system, *but also with a whole set of methodologies and procedures for gathering proper data, creating and regularly maintaining a corresponding information system suitable for supporting decision making in transport policy matters*.
- In the case of ETIS, the usual top-down design method, used for the national systems cannot work. Instead, a combination of this approach with a bottom-up one has to be used:

  Top-down part of the design of ETIS:

  Policy issues → Formulation of the questions, to which the IS has to provide answers→ Initial decisions on the kind of desired data and algorithms that have to be included in the system

  Bottom-up part of the design of ETIS:

  Determination and harmonisation of the *available* data or the data that *can be* collected → Given these data, determination of the possible algorithms and processing methods that can be successfully used → Identification of the questions, to which the resulting information system can answer → Building of the system
- It may be necessary to build specific, *localised ETIS sub-systems* for dealing with problems and political issues of multi-national character, which are particular for certain areas. These systems can be though as a more complete and extended form of the already introduced transport observatories. An example of such a system is the Alpine Transport Information System (ATIS), studied in the Transalpine Pilot Study in MESUDEMO.

Similarly, *there may be ETIS sub-systems that will be used for transport planning at the European level and are covering all Europe, or at least a large part of it*; or sub-systems that will be used for studying of *European-wide* situations or problems.

- Therefore, a possible and, actually, probable form that an ETIS can take is that of *a network of interconnected, co-operating (i.e. federated) systems*: **the *ETIS network***. The transport experts must rather think of an ETIS network, or its member systems (localised or European-wide), than of a single system with all the information in it.
- One must carefully consider not only the *initial creation* of ETIS, but also the *maintenance and update* of it as major challenge. In fact, building a complete information system of this type is only justified, if this system is *highly re-usable and updateable*.

On the overall architecture of an ETIS

- The following statement refers to any ETIS sub-system (i.e., any member of the ETIS network). In order to:
  - Cope with the very increased complexity of the procedure for the creation of such a system,
  - allow the proper maintenance and updating of this system, and
  - allow this system to co-operate to the other ETIS sub-systems and be properly *federated* into the ETIS network,

  it is strongly recommended to consider an overall architecture like that, which is briefly described in this report.
- As far as the technical design of a member of the ETIS network is concerned: Instead of simply adding a user interface and a GIS as complements to the data core and the models of ETIS, it strongly recommended to follow an up-to-date design method and *use a modern, 3-tier architecture for ETIS, which technically guarantees openness and scalability of the system and facilitates the re-usability of the data core of ETIS*. Additionally, it is recommended to also employ *Web-enabled interfaces,* which will considerably enhance the access to ETIS over the Internet.
- Regarding the data core of an ETIS (or the data, which may be distributed over the ETIS network, in the data cores of the federated members of it): *Data interoperability* is one of the most crucial points in the effort of achieving an ETIS. Standardisation may provide an answer to the data interoperability needs. However, at this moment, *it seems highly improbable to achieve a single data description and specification, covering the necessary data for working on all the problems that an ETIS (or the ETIS network) may have to deal with*. Therefore,
  - *Any standardisation effort must always refer to the specific problems, for which the data are needed,* since this is the way to choose in a proper way standard forms and methodologies for creating primary data sets (standardisation is actually *issue-relevant*).
  - Since standardisation does not always provide a total solution to the data heterogeneity problem, *mediation must be considered as an alternative to standardisation* (mediation refers to the creation of a module for translating data between two systems with different data descriptions and internal representations).

  In the cases, in which data cannot be made compatible or interoperable, therefore mediation is impossible, *the only solution is to revert to the actual sources of the data and encourage and promote full compatibility at the data collection level*.

On the internal data architecture of the data core of an ETIS (the data models for the networks, the flows and any associated information)

- In any ETIS:
    - *Consider all the categories of data that may be present in the system, as they are described in this report* (sub-section "The different kinds of data in ETIS – Ways to organise them" of the section 6.6).
    - Moreover, *always consider the meta-data, which must accompany the data sets*.
    - *Organise these data in a modern DBMS (Database Management System), and not a GIS*. Avoid the usual "planar" organisation of the data, which is customary for the GIS-centric organisations. Instead, adopt an organisation method, which at any given moment can *dynamically produce* a "plane of visible information" (the *Visible Plane*) for viewing purposes or for further computations, *by combining and mixing information at the same or different Levels-of-Detail*. (Please see the paragraph "The different kinds of data in ETIS – Ways to organise them" of section 6.6 in this report). This way, a *variable granularity of the information* can be achieved; such a variable granularity will be needed in a lot of "localised" (regional) studies (such as the Transalpine Pilot Study).
    - The form and the internal representation of the data of an ETIS (actually of any IS) absolutely depend on the issues (policy, transport, socio-economic issues or options, or any problem), on which an answer has to be provided. Moreover, in the effort to re-use the data of ETIS, by using the same system (or network) for answering to or studying more problems, a considerable diversity and polymorphy of data may result. Therefore:
      Use a special internal architecture for the data core of any ETIS, in order to anticipate the diversity and polymorphy of the data; *clearly, the mere inclusion of meta data is not enough*. A novel way to handle this problem is the proposed in MESUDEMO: *the issue-relevant organisation of the data*, in which, *not only the data themselves, but the very issues or problems, for the tackling of which the data will be used, must be stored in the IS and used for a kind of "conceptual sorting" of the data*.

      By following this methodology, highly diverse and polymorph data sets can be easily accommodated in the same system (or network), while customised views of these data sets, highly tuned to the issues or problems at hand, can be provided to (i.e. easily and transparently selected for) the ETIS users.
    - Organise the main ETIS data with the *infrastructure networks* and the *regions* as *reference information.*
- As far as the <u>*network information*</u> is concerned:
    - *A <u>unique</u>, <u>multi-modal</u>, <u>multi-scale</u>, <u>multi-purpose</u> network must be used for the European network database*. The different scales are needed for tackling problems, which require network-related information at different Levels-of-Detail (LoD). Three scales for LoD are considered enough.
    - *A permanent solution for the creation, maintenance and possible future expansion of the European network database must be adopted*. This solution must be based on the acquisition of the corresponding data from these sources, which create and maintain the corresponding primary information; namely, the Member States. Therefore:
      Each Member State should provide to the Commission network information for its area, according to a methodology, which has to be agreed upon beforehand. Then the Commission can put the partial information together, and resolve, if necessary, any arising problems in unifying the information in co-operation with the Member States.

  - It is quite probable that *several networks* (reference networks at various scales, special networks produced for specific studies, several modelling networks) *may co-exist in ETIS. The internal architecture should be able to cope with this demand.*
  - *The internal architecture of the networks should be flexible enough to satisfy emerging or even future needs*: Special sub-networks for tackling emerging problems (e.g. the movement of special new freight carrying units, combined and intermodal transport, the use of new transport models that require specific *modelling networks*, etc.) should be created relatively easy, with a considerable degree of automation and with significantly reduced cost.
- As far as the *internal data architecture* of the networks, the flows and any information associated to them is concerned:
  - *An internal architecture, employing the principle of an issue-relevant organisation and allowing multiple sub-networks and sub-regioning schemes and information at various Levels-of-Detail must be used in the heart of the data core of ETIS.*

    The basic lines of such an internal architecture are briefly presented in this report (section "A novel, unified, internal data architecture and the corresponding data model for the infrastructure networks and the flows").

  - In the internal data architecture, an *elemental representation*, i.e. a very simple, *code-based* representation, *without any geo-referencing or any geometry,* of both the networks and the regions should be employed in the database. Any other information, including geo-referencing or geometry, must be linked to this elemental representation as *attributes* or *key-linked, associated information*.

    In order to achieve this and *create pan-European infrastructure network databases*, *a proper nomenclature, i.e. a hierarchical coding scheme for clearly and unambiguously identifying each infrastructure network link and node, from the EU level down to the national level, must be created.* A corresponding nomenclature for the regions already exists: The NUTS zoning scheme.
  - As already mentioned, *the infrastructure networks* and *the regions* must be the *reference information* for any regional or European-wide associated information. In that sense, the infrastructure networks and the regions act as *integrators of the information*.
  - Regarding the infrastructure networks: *A common referential system for the geo-referencing should be employed*. Moreover, *the Commission and the Member States must agree on a unique key-system for all Europe for linking information to the networks*.
  - *Routes and services* should also be represented in ETIS, as described in this report ("The data model for the networks, the regions and any associated information").
  - As already mentioned, *all data should be organised in a modern DBMS, in which any geo-referencing should be included*. A *Spatial Data Interchange Module* should be employed for linking this database to a GIS for visualisation purposes and spatial analyses (section "The data model for the networks, the regions and any associated information" of this report).
  - *Stamping the data* with special stamps, such as a *time stamp*, a *Level-of-Detail stamp*, a *confidentiality stamp*, a *data quality and certification stamp*, etc., will greatly enhance the operation of the data core (again, section "The data model for the networks, the regions and any associated information" of this report).

On the technical construction of ETIS - Methodology and procedures for integrating (or working with) external data sets, models and sources

- *Instead of creating a closed, proprietary system, create an open configuration, which can serve both as a model for other systems of this kind, and as a "platform" for attracting the creation of third-party "add-on" programmes and the integration of third-party data sets and processing methods (including models).*

  This mentality is actually in line with the current international trend of *sharing expensive geo-information*, and at the same time, *helps in implicitly achieving a broader data and processing methods interoperability*.

- An *open and portable development (programming) environment* has to be used for the realisation of ETIS (e.g. a Java-based one, due to its open character, its portability across hardware and software platforms and its low cost; other commercial alternatives are possible, as well).
- A methodology for *the seamless and technically sound integration of third-party data sets and processing methods (including models) in ETIS*, which *guarantees honouring of any existing confidentiality requirements*, must be employed in ETIS.

General technical recommendations for ETIS

- In general: *Carefully avoid building restrictions in any ETIS system. If the system has to be restricted, due to budgetary considerations, restrict the data set and not the architectural design, so that the possibility of expanding the data set later is kept open.*
- Moreover, *follow a standard software development methodology, as developed in the field of software engineering; e.g. the standard ISO 9000-3 for guaranteeing the quality of the software production procedures, or the more general TickIt for certifying the quality of the software products*.

6.10 The MESUDEMO demonstration and dissemination tools

In order to achieve an effective dissemination of the results of the work on the ETIS overall and internal architecture, the MESUDEMO consortium has constructed a tool for the presentation of the new concepts, ideas and methodologies for the creation of an ETIS and the demonstration of the advantages of the novel points of the proposed architecture.

This tool is a Web-enabled environment, with a 3-tier, client-server architecture and advanced capabilities for handling geo-information through the Web. Web-based hypertext (html documents) will be used for the presentations, so that the users (i.e. the target audience for the dissemination and demonstration) will be able to concentrate on the points of their interest. Special clickable buttons in the Web documents will initiate demonstration programs.

The same environment can be used for the dissemination of the results of the Transalpine Pilot and the rest of the MESUDEMO Work Packages.

A PC with MS Windows and an Internet connection will be enough for accessing the demonstration and dissemination tools. Any user will be able to download through the Web and install on his PC the client part of the environment (i.e. these programs and data, which will re-side on the PCs of the users). This part will communicate with a Web-based Server in NTUA (Athens), in order to provide the presentation and the demonstration.

The demonstration and dissemination environment will be accessible through the MESUDEMO Web Server, the address of which is:
http://www.telecom.ntua.gr/MESUDEMO

7 OVERALL CONCLUSIONS AND RECOMMENDATIONS

Project setting
The implementation of the Single Market in the European Union (EU) from 1993 and the abolishment of customs barriers in the EU has led to both a decline of transport data availability and a deterioration of the quality of existing data on border crossing transport. At the same time as demand for more and better statistical information on international transport is growing, supply is deteriorating.

Suitable tools and methods are lacking on the international level to process existing data into information that satisfies the needs expressed in the Common Transport Policy (CTP) of the EU. In particular methods for the quantification of traffic and transport demand on say NUTS levels 2 or 3 are not generally available at these levels. Although methods for traffic and transport analysis are becoming ever more refined on the national level, these tools are hard to exploit on the European level because of the poor availability of necessary empirical data. Efforts are currently being made through the 4th and the 5th Framework Programmes of the EU to change and reverse this deplorable trend.

The MESUDEMO consortium was given the task to develop methodology for compiling, storing, maintaining and disseminating information on transport networks and on flows of goods and passengers. An integral part of the task was to develop data architecture and data models adapted to these flows and networks. The European Transport Policy Information System known as ETIS and developed in the INFOSTAT project has been the foundation of the work in MESUDEMO.

The quest for statistical information to guide transport policy decision making can be viewed as a top-down identification process of necessary data. On the other hand various kinds of transport statistics have existed for years. So the question might be formulated as what kind of transport policy issues can be viewed in the light of existing data. This is the bottom-up process of identifying the political relevance of existing data. During the project work the consortium has all the time consciously striven to find some workable compromise between the two approaches which in the end will determine the degree of successful implementation of ETIS.

The challenge of promoting knowledge about the efficiency of transport infrastructure utilisation resulted in an organisation of MESUDEMO along two lines. One line treats data on infrastructure provision and the other deals with data on transport flows constituting demand for transport infrastructure. Both lines are joined together through the use of state-of-art information technology adapted to the needs of MESUDEMO. This has been a major achievement of the project.

MESUDEMO was originally designed as parallel strings of work packages. However, in the course of work it was experienced that later work packages had impacts on previous ones, spurring further work there. This in turn laid the foundation for more work in the following work packages. Thus experience necessitated a dynamic way of working and resulted in most of the Deliverables being completed at the end of the project.

The MESUDEMO project has employed a broad set of research methods, varying from literature studies, questionnaire surveys, in-depth interviews, pilot studies and workshops.

Transport flows

The current situation as to data availability is rather bleak. Data on European goods transport flows are mostly restricted to a unimodal registration at an annual basis of tonnes and/or tonnekilometres broken down by commodity groups and only specifying country of origin (O) and country of destination (D) (NUTS 0). Current data on transport chains (intermodal transport) do not as a rule exist. For passenger transport even much less data are available. As an overall conclusion it can be maintained that there are very few O/D related data available for border crossing passenger transport.

It is recommended that ETIS should be based on a transport policy oriented, comprehensive database with harmonised data elements. MESUDEMO recommends that a European transport policy oriented database should be based on the INFOSTAT list of indicators/ variables and focus on O/D flows. The ambition might be O/D flows at NUTS 3 level for both goods and passenger flows, since this is a level at which Pan-European transport policy may be reasonably well monitored and corridor studies evaluated. From NUTS 3 the data can be aggregated to higher levels and thus be used to supply answers to Pan-European policy questions at more aggregate levels. However, considering the current situation on data availability, the NUTS 2 level might be the most ambitious to aspire to.

The NUTS zoning is based on population figures. A zone on the NUTS 2 level in a densely populated country such as the Netherlands or Belgium is much smaller in area than a zone on the same NUTS level in the northern part of Sweden or Norway. This is not the best situation for traffic and transportation studies. Particular problems are incurred in the selection of the different network elements. Still at the European level, data at a NUTS 2 level may be considered satisfactory. In the case of localised observatories or policy support systems, it may be necessary to use a variable granularity of information, e.g. NUTS 3 for the region of interest, NUTS 2 in the vicinity of it and even less than this – such as NUTS 1 or 0 in distant areas.

Creating reasonable transport databases for the European policy level will be a demading task. In the ideal case the following aspects of a **goods** transport database should be kept in mind when aggregating the necessary linked transport information:

- A harmonised classification system should be used to be able to keep hold of the most detailed information possible, such as information on the load (consignment)
- Individual items should be considered as new variables, such as number of trips, number of consignments, etc
- Some additional variables should be calculated, e.g. weight x consignment distance for tonne-km, capacity x (loaded and/or empty) trip distance, etc
- Consistency with classification used in trade databases for building transport chains should be provided
- Consistency with classification used in other parts of ETIS (socio-economic data, passenger flows, infrastructure links and nodes, impacts) should also be secured.

One main achievement of the Transalpine Pilot Study has been to demonstrate that trade and transport databases can be successfully harmonised, through the construction of a prototype demonstrator tool for studying policy options in the Alpine Area. The pilot study has demonstrated that it is possible to aggregate and merge information from different, international and national sources, in order to build a new, harmonised database for policy support.

In the pilot study called TRANSITIE an inventory of statistical problems was made and a methodology was successfully tested and proposed to overcome the problems related to the establishment of a transport chain database. Dutch databases were used in this pilot study. However, considering the difficulties experienced in the Netherlands, and as it cannot be ruled out that similar data problems also exist in other European countries, the creation of a Pan-European consistent transport chain database will be rather challenging and complicated.

Concerning **passenger** transport, international comparison and harmonisation is generally difficult due to various zoning systems, sample figures, modelling procedures, definitions on transport modes etc. A great effort is consequently needed to get the different countries to enrich and adapt their present system of data collection.

The opportunity should be taken to harmonise the definition of concepts and nomenclature of variables within each country and among European countries, in order to reach a consistent set of surveys at the European level. Communication and co-operation with the national institutions will in this respect be of central importance in order to succeed in building ETIS as a transparent and multi-layered information system.

By combining a top-down and bottom-up approach, and gradually building and supplementing the system with additionally collected data, a passenger component of ETIS can be developed over time, leading to a more reliable modelling of O/D matrices based on a better coverage through observed data. A group of national public road administrations have laid the foundation for an information system for the European road network (the RADEF project). The system includes geo-referenced data on traffic accidents, traffic counts, etc. The next recommended step would be to attach also O/D data to the network.

Transport infrastructure networks

The MESUDEMO consortium has made a comprehensive survey of available information on transport infrastructure networks and attributes linked to them. The consortium has developed two possible solutions for the data architecture of the networks. The first one is to make use of the network as an integrator of data, and to let a Geographical Information System (GIS)be the backbone of the database. The other approach is to take advantage of the fact that information in ETIS is necessarily georeferenced and can be organised according to region or the network of the transport links and modes. A unified model has been developed for ETIS, for handling all these different kinds of data.As a general rule, it is recommended that all data should be organised in a contemporary, advanced Database Management System (DMBS), instead of a GIS.

Road networks

The only way to guarantee an up to date accurate road network with many attributes is by the utilisation of national sources. The national road networks are maintained in most countries on a continuous basis. Generally, it is recommended to:

- define the ETIS-relevant roads;
- establish uniform definitions of attributes;
- establish an international organisation to receive the data from the national organisations on a continuous base.

A first step to the above-mentioned topics is established by the RADEF project. The RADEF data dictionary provides agreed terms, definitions, categorisations, and relationships for various types of road data, such as traffic details, number of accidents, road geometry and road condition statistics. Common location referencing is a particularly important component, linking the national networks to a European network. The data dictionary is the result of extensive co-operative research by the public road administrations. The RADEF project may be adjusted and further developed through co-operation for the purpose of creating the first ETIS network for roads. It will be a task for the Commission to encourage further co-operation with the national road administrations.

Networks for other modes
Changes in the infrastructure for other modes than roads normally occur less frequently. The recommendation is to continue the maintenance of existing networks on an ad-hoc base. It is also recommended to establish an organisation to guarantee the quality of the network data. Furthermore, a standardisation of the definition of attributes is necessary.

For assignment of transport chains, an intermodal network is essential. Apart from the ETIS-relevant attributes a unique transshipmentipment location must be added also to relate information on to the separate networks.

Compatibility issues

The recommendation is to harmonise and to make existing data and systems interoperable, and to continue to encourage an explicit conceptualising and theory-building process. In this manner we will not only understand mobility better, but we will also understand it in a similar manner around Europe. The research projects in the Framework Programmes, of course, have all been related to this European level, with partners from the various countries trying to grasp European transport development. MESUDEMO participated and contributed in this dialogue through two workshops, in-depth interviews, questionnaire survey, many meetings with external participants and general dissemination of the work.

Implementation and use of ETIS will make conceptualisation explicit. The recommended architecture of ETIS allows the system to function as a mediator in a dialogue, none of the modules in the architecture are thought to be automatic and static, but should rather be considered as arenas for the converging and learning process towards understanding mobility.

Data architecture and data models

The main conclusions from the work on data architecture and data models can be summarised as follows:

- A better understanding of the organisational and technical nature of ETIS has been achieved and the corresponding basic concepts have been described.
- A general architecture of ETIS has been developed, and its main modules have been analysed further.
- A novel internal data architecture and the corresponding data model for the infrastructure networks and the flows have been developed; this internal architecture has enhanced capabilities of handling transport data.
- A realisation methodology for ETIS, which reduces the installation and maintenance costs and permits the open (but controlled) access to its data and processing methods, has been

developed. This methodology, moreover, permits the easy construction of third-party additional software for ETIS.
- A concrete need for preparatory work on behalf of the Commission at the European level, in order to achieve the ETIS deployment, has been identified.
- The Transalpine Pilot Study has brought forward a demonstrator showing the novelties of the proposed architecture
- Moreover, a step-wise, practical path to a future ETIS has been envisaged and described.

The following figure outlines the general architecture of ETIS, showing the various sub-systems at a conceptual level.

A general architecture of ETIS, showing the various sub-systems at a conceptual level

ETIS will comprise data of two categories; structural and unstructural data. The first category can be further divided:

- *Static data*, i.e. data that regularly do not change or change very slowly. Usually most of the infrastructure-related data as well as socio-economic data fall in this category.
- *Dynamic data*, i.e. data that change very fast and need special methods for handling, e.g., traffic data on a road network that may need updating very often.
- *Time series* comprising the time pattern of indicators.

Unstructured data are usually textual or multimedia data, including documentation, EU or national directives, knowledge representation, etc.

Usually, a Geographical Information System (GIS) is used for the organisation of the infrastructure-related data. The transport networks, specifically the nodes and the links, are used as the elements, to which all other information is linked. Each mode is usually on a different plane, called layer. The work in MESUDEMO, however, has shown that such a conventional organisation can prove both inefficient and restricting for ETIS.

In MESUDEMO, a unified model has been developed for ETIS, for handling all different kinds of data, as well as the necessary meta-information. As a general rule, all data are recommended to be organised in a contemporary, advanced DBMS (Database Management System). The need for geographical visualisations and spatial analyses are covered through the use of a spatial data interchange module, which takes care of the communication of the database to a GIS. Such modules have already become commercially available.

ETIS must be associated not only with an information system, but also with a whole set of methodologies and procedures for gathering proper data, creating and regularly maintaining a corresponding information system suitable for supporting decision making in transport policy matters.

In the case of ETIS, the usual top-down design method, used for national systems cannot stand alone. Instead, a combination of this approach with a bottom-up one has to be used. It may be necessary to build specific, localised ETIS sub-systems for dealing with problems and political issues of multi-national character particular for certain areas. These systems can be thought as a more complete and extended form of transport observatories. A possible and probable form that an ETIS can take is that of a network of interconnected, co-operating (i.e. federated) systems: **the ETIS network**.

One must carefully consider not only the initial creation of ETIS, but also the maintenance and update of it as major challenge. In fact, building a complete information system of this type is only justified, if this system is highly re-usable and updateable.

In the process of creating an ETIS, it is very important to carefully avoid building restrictions into the system. If the system has to be restricted, e.g. due to budgetary considerations, restrictions have to be put on the data set and not on to the architectural design, thus keeping open the possibility of expanding the data set later on.

A pathway for the creation of ETIS

On the preparation of ETIS the consortium has recommended not only to concentrate on questions of methodology and content but also to consider institutionalisation of ETIS. Some relevant issues to look into are:

- Which institution will be responsible for building and maintaining ETIS and what resources will be available?
- What will be the role of the European countries within ETIS? Will it be possible to convince them of the benefits of such a system?
- What regulations will have to be adapted to meet the demands of ETIS?
- How does one solve the confidentiality problem?

A stepwise, iterative, procedure has been sketched for the systematic analysis of the data, the processing methods, the architecture and the interfaces of an ETIS.

Step1: List the most pressing transport or policy problems, for which solutions should be provided by ETIS. Include typical questions, for which an answer preferably would be provided by ETIS.

Step 2: Convert these problems or questions into transport or policy options (or measures), which have to be examined as possible ways of solving the problems or answering the questions.

Step 3: List the set of the indicators, which are necessary for evaluating these options or the implications of these measures.

Step 4: Determine the methodologies, which are needed for the computation of these indicators. Determine, therefore, exactly which data and computation methods (including models) are needed. Determine also the desired Level-of-Detail and geographical coverage of the data.

Step 5: Complement the lists with data and computing methodologies (including the models) that are not indicated by the procedure so far, but may be proven useful in the future, as far as the understanding of the transport problems in the area are concerned.

Step 6: Then, go shopping for data, by contacting all those data providers, who may have the data needed. Prefer, if possible, those providers, who are responsible for the production and update of the primary sets. Determine the availability of the data needed and the way to establish the paths for acquiring and regular updating of the data. If there are any confidentiality issues regarding these data, look for an arrangement for sorting out this matter.

Step 7: If the data exist, determine what kind of possible harmonisation, validation and/or merging is needed, and go to step 8. If however, there are significant gaps in the data, or the data would not be disclosed, look for alternative sources; examine if these sources can provide data at the desired quality and regularity. If proper sources are found, then go to step 8. In the case there are still severe data gaps, examine if alternative ways of getting the needed data may be found, such as possible reliable estimation of the missing information, or possible collection of primary data, etc. If such alternatives arise, and applicable ways can be found, continue to step 8. Otherwise, examine if desirable results can be produced with the (not complete) data set that may be acquired, and if it is still worthwhile to create this ETIS at this moment. If "yes", adjust (i.e. reduce) the set of transport problems or policy issues that can be answered, execute steps 2, 3, 4 and 5 again and then go to step 8. If not, drop the job for the time being. There is always a chance that the missing data will be available in the future through collection of primary data or otherwise.

Step 8: Having determined the questions, to which the system can provide answers, given the data availability, technically design the system at the conceptual level, following a top-down approach. The process described in the previous steps would have already formed the Issue-relevance Table to be used in organising the data of the system.

Step 9: Decide on the software platform, the development environment and the software tools, which have to be used. Use a platform, a development environment and software tools, which guarantee an open and scalable architecture, and encourage the construction of third-party software.

Step 10: Based on the decisions for the software platform, the development environment and the software tools, as well as on adequate samples of the data, conclude the design (logical and physical) and the construction, first of the basic data-centric modules and then of the rest of the modules of the overall architecture. Conclude the design (logical and physical) and the construction of the basic data-centric modules first, based on the decisions for the software platform, the development environment and the software tools, as well as on adequate samples of the data. And only afterwards conclude the design of the rest of the modules in the overall architecture.

Step 11: Based on the decisions taken in steps 6 and 7, acquire the data. Some of them should already be available from a permanent, centralised acquisition activity; e.g. the infrastructure networks, which have to be taken from participating countries and updated regularly.

Step 12: Perform any validation, harmonisation and merging of the data, as necessary. Then, organise the data in the data-centric modules of the system. Carefully determine if there is a need for accessing remote data on demand. If so, create the necessary procedures. Moreover, determine the confidentiality resolution procedures, and program the corresponding components into the system. At the same time, organise the procedures for updating the data (since this operation may actually influence the overall data organisation). If necessary, make any required adjustment to the structure of the system components, which are used for the realisation of the modules of ETIS.

Step 13: Systematically acquire or create the meta-information, which will be necessary for the proper operation of the system. Moreover, determine and acquire the knowledge, which has to be put into the system in order to support the operation of the system and the decision making process. Organise the meta-information and the necessary knowledge in the system and link it to the data and the processing methods.

Step 14: Complete the construction of the system components, integrate the system and perform tests for assuring its proper operation and evaluating its performance. Make any corrections that may be necessary.

Step 15: If necessary, proceed to the actions needed for connecting (integrating or federating) the system to the ETIS network. The kind of connection would highly depend on the exact aim of the individual system under construction and the general ETIS needs for accessing data and processing methods.

Step 16: Carefully document any decisions and steps taken for the design and construction of the whole system, starting from the decisions on the issues and problems that the system should deal with and the data it would contain. This is a vital step for the proper, continual operation of the system, as well as for any update or expansion of it.

MESUDEMO has been a large project, in resources deployed, in time-span, and in thematic scope. The project underpins the need for continuity in transport research compatible with the continuing and ever-changing character of society itself. An enhanced understanding of transport dynamics is needed to guide policy design, especially in the long-term perspective. In the short run, existing tools (as transport models and cost/benefit analyses) can give relevant answers to pressing problems, as long as adequate data are provided. Suggestions on how to extract, process, manage, and disseminate such information at the European level through the use of state-of-the-art information technology, are major achievements of the project. In the long run, however, actual transport developments will set the agenda for raising new policy issues. Such long-term transport development is not easily calculated on the basis of current appearances, but can at least be better understood by disclosing the driving forces of the system. By enabling this, we may increase our ability to design policy measures in due time, and enable us to balance different measures against each other.

LIST OF ABBREVIATIONS

AAD	Aan-, af- en doorvoer (Import, Export, Transhipment)
ADT	Annual (Average) Daily Traffic
AF	Agder Research Foundation
AGC	European Agreement on Main International Railway Lines (UN-ECE)
AIC	ArcInfo Coverage
ANWB	Dutch Automobile Association
APAS	Actions de Préparation d'Accompagnement et du Suivi
APUR	Atelier Parisien d'Urbanisme
AQUABEL	Incident report (GIS) system of the Dutch Ministry of Transport
ArcInfo	GIS software by ESRI
ASECAP	European Association of Toll Motorway Concessionaires and Agencies
ASSEMBLING	4th Framework Project: Assembling a European network of monitoring centers for transport infrastructure
ATIS	Alpine Transport Information System for decision support: a localised ETIS (see corresponding entry).
ATS	American Travel Survey
AVV	Adviesdienst Verkeer en Vervoer (see TRC)
BRTN	Beleidsvisie Recreatie-toervaart Nederland policy)
CBS	Dutch National Bureau of Statistics
CBS	Copenhagen Business School
CEEC	Central and Eastern European Countries
CEMT	See ECMT
CETMO	Public foundation of Spanish transport
CIS	Commonwealth of Independent States
Conceptual level of the design of an Information System	The design of the whole Information System at a functional level, independently from the software and hardware, which will be used
CONCERTO	4th Framework Project: Concerted action for European transport information systems
COROP	Dutch nomenclature comparable to NUTS level 3
COST	Coopération Européenne dans le domaine de la recherche Scientifique et Technique
CTP	Common Transport Policy
CTSE	Commodity Classification for Transport Statistics in Europe
CWR	Continuous Welded Rail (rail construction system)
Data Model vs. transport Model	"Data model" is an information technology term and signifies the internal data structure of an information system. The term "Data model" is completely irrelevant to the well-known to the transport experts term "transport model".
DATELINE	5th Framework Project

DBMS	Database Management System
DG TREN	Directorate-General for Energy and Transport of the European Union
DHTML	Dynamic Hyper Text Markup Language. DHTML combines HTML, style sheets, and scripts to make Web pages more interactive and permits an elementary computational ability
DLL	Dynamic Link Library
DRIVE-	Dedicated Road Infrastructure for Vehicle safety in Europe European Community
DSS	Decision Support System
Dynamic segmentation	Technique to store, query and analyse the attributes of parts of a linear feature without having to modify the whole network feature.
EC	European Commission
ECDIS	Electronic Chart Data and Information System
ECE	See UN-ECE
ECMT	European Conference of Ministers of Transport
ECU	European Currency Unit
EDI	Electronic Data Interchange
EEA	European Economic Area
EIB	European Investment Bank
EIMP	European Infrastructure Master Plan
ERTMS	European Rail Traffic Management System
ESRI	Producer of ArcInfo
ETIS	European Transport Policy Information System
ETIS network	A network of ETIS-type systems, comprising any localised sub-systems (or observatories) working at a regional, but multi-national level, as well any European-level (supra-national) sub-systems. These systems should be compatible in the way they are handling their data and processing methods (including models) and *federated*, in order to provide uniform access to them.
EU	European Union
EURET	European Research Programme for Transport (from 2nd Framework Programme
EUROSTAT	Statistical Office of the European Communities
FYROM	Former Republic Of Macedonia
GDF	Geographic Data File
GEOSYSTRANS	4th Framework project: Geometic database on European transport
GIS	Geographical Information System
GISCO	GIS system of EUROSTAT
GMS	Geïntegreerd Meldkamer Systeem (Dutch joint incidentroom system)
GN	Measurement unit per individual goods type of the harmonised system

GNC	Global Navigation and planning Charts
GTF	General Transportation data Format
HST	High Speed Train
HTML	Hypertext Mark-up Language: It is the hypertext (see relative entry) “language” for the Web (which includes marks for the formatting of the text).
Hypertext	A text format, in which links to other pieces of text of the same or other documents can be included; use of these links brings up directly the linked text. An example of this is the “HTML” format that the Web uses.
IATA	International Air Transport Association
ICAO	Intarnational Civil Aviation Organization
ID-	Identification tag
IMEG	The Eurostat Intermodal Expert Group
INDRIS	Localisation and tracing system for the Dutch NIN
INFOSTAT	4th Framework project: Information Systems and Statistics Strategic Transport
INRETS	The French National Institute for Transport and Safety Research
INTRASTAT	Intra EU Trade Value Statistics
IQ	A transport RTD Programme for intermodal transport
IRPUD	Institut für Raumplanung, Universität Dortmund
IS	Information System
ISO	International Organization for Standardization
Issue-relevance	The way of organising diverse and polymorph information according to its relevance to the main issues, to which an Information System is called to provide answers (in the case of ETIS, transport or Policy issues). This term is introduced in this document for the first time.
ISTAT	National Statistical Office of Italy
IT-	Information Technology
ITEP	Institut des Transports et de Planification
IWG-	Intersecretariat Working Group. Forum for multilateral co-operation in transport statistics. Members are UN/ECE, Eurostat and ECMT
LAN-	Local Area Network
Layer (in a GIS)	A “plane” of a conventional GIS for showing thematically similar data, which usually are at the same Level-of-Detail. E.g., A plane with the roads, another with the rail, another with the regions, etc.
Localised ETIS	An ETIS used for *local planning*, i.e. for working on problems specific to a *multi-national, but restricted* (i.e. not pan-European) area (see ch.2.2.2)
Logical level of the design of an Information system	The design of the whole Information system with a given data modelling technique, independently from the actual software packages and hardware, which will be used.

Logical vs. Physical sub-set of data	A physical sub-set of data is meant as a separate dataset on a storage medium; a logical sub-set is one that actually lies in a larger, organised set of data and is distinguished from the rest of the data either by using specific pointers, which are pointing to these data, or by providing a proper selection method. The logical sub-set can easily be extracted and stored on a medium, or organised in a separate database; in such a case, it becomes physical.
LOS	Level Of Service
MDS	MDS Transmodal Ltd.
ME&P	Marcial Echenique & Partners Ltd.
MEST	4th Framework Project: Methods for European Surveys of Travel Behaviour.
MET	Ministère Wallon de l'Équipement et des Transports
Meta-Information	Information about some other information (usually the data of a system). Meta-information usually describes the meaning and structure of this other information.
Meta-Information Server	The server (programme or computer), which provides meta-information
MGE	Producer of software
Mkmetric	Transport consultant
Module-External Meta-data	The meta-data, which refer to the data of a module, but are stored *outside* this module (e.g. in a knowledge base)
Module-Internal Meta-data	The meta-data, which refer to the data of a module, and are stored *inside* this module (e.g. in an internal sub-module for handling of this information)
Multi-scale networks	Networks containing information in several Levels-of-Detail (LoDs). Special architecture are necessary for organising, storing and inter-relating the information in such networks.
MVA	Transport consultant
MYSTIC	4th Framework project: The Methodology and Evaluation Framework for Modelling Passengers and Goods on Transport Infrastructure Scenarios using the ETIS
NEA	NEA Transport Research and Training
NEAC Model	Transport model developed by NEA
NEN	Dutch standardization institute
NIN	Dutch National Infrastructure Network
NST	Nomenclature Uniforme des Marchandises pour les Statistiques de Transport
NSTR	Nomenclature Uniforme des Marchandises pour les Statistiques de Transport Revisée
NTUA	National Technical University of Athens
NUTS	Nomenclature of Territorial Units for Statistics
OD	Origin Destination
OD-ESTIM	4th Framework Project: Cost-efficient Origin/Destination Estimator
ONC	Operational Navigation Charts

O-O	Object-oriented (technology, programming, etc.)
ORACLE	Database software
Partial attributes (mainly related to the links and nodes of the transport networks)	Attributes which can be assigned to a part only of a link of a network
PAWN	Policy Analysis Waterway management for the Netherlands
PC-Navigo	Inland navigation software
PCU	Passenger Car Unit
PHARE	EU's programme for financial and technical co-operation with CEEC counttries:Albania, Bosnia and Herzegovina, Bulgaria, Czech Republic, Estonia, Former Yugoslavian Republic of Macedonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia
Physical level of the design of an Information system	The complete design of the Information system, down to the individual variables and programming sub-modules
POF-name	Part Of Fairway name
Polymorph data	Data, for which more than one form and format, are allowed in a system. In order to use these data, the user must explicitly or implicitly (through the *issue-relevance*, see the corresponding entry in the glossary) select a proper form and format for them.
Polymorphy	The existence of *polymorph data* in a system (see the corresponding entry in the glossary). Special techniques are usually needed for handling polymorphy.
POP-name/-number	Part Of Port name/number
PPP	Public private partnership
PRA	Public Road Administration
PTT	Dutch Mail
RADEF	Road Administration Data Exchange Format
RDBMS	Relational Database Management System
Realisation of an Information system	The actual construction of the Information system.
RIV	UIC's regulation scheme of freight cars
RIZA	Dutch instutute for fresh and wasted water management
RTD	Research and Technical Development
SAD	Single Administrative Document
SAGITTA	Dutch Computer System for Customs Declarations
SBA	Statistisch Bundesamt, German National Bureau of Statistics
SCENES	4th Framework project: European Transport Scenarios
SITC	Standard International Trade Classification
SITRAM	Système d'Information sur les Transports de Marchandises
SLIM	Supply Chain Logistics and Information Management
SOFRES	Taylor Nelson Sofres Consulting
Specific views of (sub)sets of	When a database contains information in many forms and

data	formats (polymorph data), it is the user needs that defines which of these forms and formats have to be used by the user. Therefore, a mechanism is needed for properly selecting the forms and formats that the user needs, creating thus a specific, logical or physical (see “logical level …” and “physical level …”entries in the glossary) “view” of the data or a sub-set of them. These data can then be exported to a new dataset or database, if needed.
STEMM Model	Model developed in the Strategic European Multi-Modal Modelling project by MDS
STREAMS Model	Model developed in the SCENES project by ME&P
Structured Data (or Information)	Data have a fixed form and format
System-oriented Meta-data	Structured information, which describes the data in a data set or a database; it includes a definition of each datum, as well as a description of its form and format. This information is essential in creating easily maintained, multi-level systems and in permitting different systems to co-operate. *The internal meta-data are provided for the use of the system, since this way certain of its operations can be further automated*. (This term is introduced in this document for the first time.)
TACIS	Similar to PHARE
TEM	“Transport-Economisch Model” (Transport Economic Model)
TEN	Trans European Networks
TEN-T	Trans-European Networks- Transport
TER	Trans European Railway
TERN	Trans European Road Network
TEU	Twenty Foot Equivalent Unit (used for containers)
TGV	Train à Grande Vitesse
TINA Ten-Ten Project	Traffic Forecast on the Ten Pan-European Transport Corridors of Helsinki by order of Phare
TIS	Transport Information System.
TRAINS	The Eurostat Transport Statistical Information System
TRAINS Database	The EUROSTAT Transport Statistical Information System
TRANSITIE	Future Registration Import, Transport and Transit the Netherlands & Statistics International Transport in Europe
TRC	Dutch transport Research Centre (see AVV)
TREX	Trade by Mode (EUROSTAT)
UIC	International Union of Railways
UN-ECE	Economic Commission of Europe of the United Nations Organisation
Unstructured Data (or Information)	The data, whose structure does not have to be necessarily known beforehand, and thus, allow the user a considerable flexibility in choosing the type and format of this information. Usually this term is contrasted to information organised in a fixed way in tables, records, or other structures used in databases.

UoD	Universe of Discourse (smallest functional unit in ETIS)
User-Oriented Meta-Information	Unstructured additional information or knowledge, related to certain data. This may include any additional documentation, regulation, comment or relative knowledge. *The external Meta-data are provided for the use of the expert user of the system*, in order to help him better understand what he can do with the data, the quality of them, how he can further use them, etc. (This term is introduced in this document for the first time.)
UTS	Producer of software
VAT	Value Added Tax
VESON	Producer of software
VHF	Very High Frequency
ViN	Vaarwegkenmerken in Nederland (Dutch inland waterways database)
Virtual database	A database, in which the dataset is actually distributed over more than one, possibly remote, computers. However, its users access the virtual database as if all the data were "local".
VAARGIS	Digital fairway network (of the Dutch TRC)
WAN	Wide Area Network
Web browser	Web browser is the interface to the World Wide Web. It can follow hypertext links and let the user visit Internet sites and navigate from one Internet machine to another.
Web-enabled user interfaces	User Interfaces that permit transactions to the systems over the World Wide Web
WP	Work Package

ANNEX

Road network attributes

List of attributes (most relevant)	GISCO	GISCO (APUR extension)	RADEF
Classification codes			
MOBILE NETWORK ELEMENTS			
TRANSPORTING UNITS (Passenger cars, truck and coach)			
TRANSPORT LOADING UNITS (seats, containers, swap body and trailer/semi – trailer)			
FIXED NETWORK ELEMENTS			ATTRIBUTES OF ROAD NETWORK DOMAIN
NODES	ROAD NODE TYPES	ditto GISCO	
Border point	BN: Border Node	ditto GISCO	e.g. border connection
Buss station	FP: Ferry Port	ditto GISCO	
Toll booth	ME: Motorway Exit	ditto GISCO	
	MI: Motorway Interchange	ditto GISCO	
	RN: Road Node	ditto GISCO	
Ferry point			
Intersection			
Junction			
Transshipment terminals: (from and to various modes)			
LINKS	ROAD SEGMENT TYPE	ditto GISCO	ATTRIBUTES OF
Roads:	D: Dual carriageway road	ditto GISCO	ROAD NETWORK DOMAIN
E-roads	DE: Dual carriageway road, European	ditto GISCO	+
Other motorways	F: Car Ferry	ditto GISCO	ATTRIBUTES OF
Toll roads	M: Motorway	ditto GISCO	STRUCTURE DOMAIN
Dual carriage highways	ME: Motorway, European	ditto GISCO	e.g. road type (motorway/national road) and structure function type (bridge/tunnel)
Highways	O: Other road	ditto GISCO	
Local road link to a terminal from another network node	OE: Other road, European	ditto GISCO	
Other roads			
Bridge			
Tunnel			
HIERARCHIES			
NODES			
Hierarchy of individual transshipment terminals: (regional, inter – regional and European)			

List of attributes (most relevant)	**GISCO**	**GISCO (APUR extension)**	**RADEF**
LINKS	ROAD SEGMENT LEVEL.	ditto GISCO	
Regional	Road segment TEN	ditto GISCO	
Inter-regional	Road segment TEN Alignment	ditto GISCO	
European	Road segment TEN priority project number	ditto GISCO	
	Road segment TEN priority project name	ditto GISCO	
		+	
		for planned TEN codes and section codes:	
		costs, investments in 1996/1997	
		total length/section length	
		project status/type,	
		current raod type /	
		future road type	
		completion year, etc.	
VARIABLES AND INDICATORS			
SPATIAL KEYS			ATTRIBUTES OF
GEO-POSITIONS			ROAD GEOMETRY DOMAIN
Geo-references of nodes and terminals	road segment code	ditto GISCO	e.g. arc point
latitude in decimal degrees	road node code	ditto GISCO	
Name of all types of nodes	road segment country code	ditto GISCO	
Road number	road segment number European 1 (first European road number)	ditto GISCO	
Network section ID	road segment number European 2 (second European road number, if existing)	ditto GISCO	
Link markers	road segment number National 1 (first national road number)	ditto GISCO	
	road segment number National 2 (second national road number)	ditto GISCO	
	road node country code 1 (ISO code of country to which the node belongs)	ditto GISCO	
	road node country code 2(ISO code of second country to which the node belongs, in case of border nodes)	ditto GISCO	
CONSTRUCTION DATA			
NODES (year of construction, method of financing, altitude)			
LINKS			
Costs of construction			

List of attributes (most relevant)	**GISCO**	**GISCO (APUR extension)**	**RADEF**
Year of construction (or year of last full renovation)			
Length of section built		total link length	
Co-ordinates of section from/to			
Altitude			ATTRIBUTES OF
Gradient			ROAD GEOMETRY DOMAIN
Surface material			e.g. grid type
Method of financing			
EU contribution if any			
FLOW SUPPORT FACILITIES			
NODES (Handling capacities of different types of transportation bodies)			
CAPACITY			
NODES (both cargo and passengers as well as containers)			
LINKS			ATTRIBUTES OF
Road link capacity		number of lanes per direction	ROAD NETWORK DOMAIN
Design road link volume			e.g. road/section road type (number of lanes)
Design speed			
RESTRICTIONS ON USE			
NODES (opening hours, passenger and/or freight			
LINKS			ATTRIBUTES OF
Max allowed axle load			RESTRICTION DOMAIN
Max length of vehicle			e.g. restriction type
Max. allowed speed, cars			(maximum axle weight,
Max. allowed speed, trucks			height, length, weight,
Max. Allowed speed, coaches			width)
Night ban on trucks			
FLOW VOLUMES			
NODES (loaded and unloaded passengers; goods and passengers in transit splitted into several transportation modes)			
LINK			ATTRIBUTES OF
Link load of trucks			TRAFFIC DOMAIN
Link load of cars		Light weight(<3.5t) ADT 1995	e.g. traffic flow type (annual daily traffic,
General link load		Light weight holiday month ADT 1995	average flow during peak period, maximum flow)
Peak hour traffic of certain road sections (identified bottlenecks)		reference months for lightweight traffic	
07.00 - 09.00			
12.00 - 13.30		Heavy weight(>3.5t) ADT 1995	
15.30 - 17.30		Heavy weight holiday month	

List of attributes (most relevant)	GISCO	GISCO (APUR extension)	RADEF
		ADT 1995	
Share of inter-regional traffic		reference months for lightweight traffic	
Share of trucks			
ENVIRONMENTAL PRESSURE			
LINKS / NODE			ATTRIBUTES OF
Severe injuries		Accident rate	ACCIDENT DOMAIN
Fatal accidents		(accidents/million vehiclekm)	e.g. injury type (fatal, minor, none, serious,
Noise emission		number of years of reference for accident rate	unknown)
Space consumption			

Rail network attributes

List of attributes (most relevant)	GISCO	GISCO (APUR extension)
Classification Codes		
MOBILE NETWORK ELEMENTS		
TRANSPORTING UNITS		
Passenger trains in different classes:		
Goods trains in different classes		
ON BOARD TRANSPORTING UNITS OF OTHER MODES		
Truck		
Passenger car		
TRANSPORT LOADING UNITS\		
Splitted into different types and different modes		
FIXED NETWORK ELEMENTS		
NODES	Railway node type.	Ditto GISCO
Railway station	BN: Border node	Ditto GISCO
Border point	BS: Border railway station	Ditto GISCO
Ferry point	RS: Railway station	Ditto GISCO
Transshipment terminals:		
Splitted into different modes and commodities		
LINKS	Full name of railway segment type:	Ditto GISCO
Bridge	BD: Branch line, double track	Ditto GISCO
Tunnel	BDE: Branch line, double track, electrified	Ditto GISCO
1 track line	BS: Branch line, single track	Ditto GISCO
2 track line	BSE: Branch line, single track, electrified	Ditto GISCO
more track line	MD: Main line, double track	Ditto GISCO
Local rail link to a terminal from another network	MDE: Main line, double	Ditto GISCO

List of attributes (most relevant)	GISCO	GISCO (APUR extension)
node	track, electrified	
Bridge	MS: Main line, single track	Ditto GISCO
Tunnel	MSE: Main line, single track, electrified	Ditto GISCO
	NG: Narrow Gauge	Ditto GISCO
	TF: Train Ferry	Ditto GISCO
HIERARCHIES		
NODES		
Hierarchy of individual transshipment terminals: (regional/ inter regional and European)		
Hierarchy of clustered transshipment terminals: (european/ inter regional)		
LINKS	Railway segment level	Ditto GISCO
Regional	Railway segment TEN	Ditto GISCO
Inter-regional	Railway segment TEN Priority Project number	Ditto GISCO
European		
VARIABLES AND INDICATORS		
GEO-POSITIONS		
Geo-references of nodes and terminals	Railway segment code	Ditto GISCO
Name of all types of nodes	Railway node code	Ditto GISCO
Line number	Railway segment country code	Ditto GISCO
Network section ID		
Link markers		
CONSTRUCTION DATA		
LINKS/NODES		
Costs of construction		
Year of construction (or year of last full renovation)		
Length of section built		length
Co-ordinates of section from/to		
Altitude		
Gradient		
Method of financing		
EU contribution if any		
FLOW SUPPORT FACILITIES		
NODES (handling of different transportation bodies)		
LINKS		
Automatic Train Control (Specifications needed?)		Attributes of line
(type of) electrical Power system :		technical description
(type of) Signalling system:		
CAPACITY		

List of attributes (most relevant)	GISCO	GISCO (APUR extension)
NODES		
Cargo handling capacity		
Terminal capacity		
Container handling capacity		
LINKS		Attributes of line
Rail link capacity		technical description
Track gauge		
Loading gauge		
Design rail link volume		
Design speed		
Viable for dangerous goods transport		
RESTRICTIONS ON USE		
NODES		
Opening hours of terminals		
Type of Terminal (passengers and/or freight)		
LINKS		Attributes of line
Max allowed axle load		technical description +
Max. allowed speed	Use of railway segment:	Ditto GISCO
Only passengers	PG: Persons and goods	Ditto GISCO
Only freight	P: Persons only	Ditto GISCO
Passengers and freight	G: Goods only	Ditto GISCO
FLOW VOLUMES		
NODES		
Volumes of goods and passengers, splitted into different types of transportation)		
LINK		Attributes of traffic
Route schedules. O/D pairs		link + attributes of
Link load: passenger trains		line traffic UIC
Link load: freight trains		
Link load, total		
Share of inter-regional traffic (Of total link traffic)		
ENVIRONMENTAL PRESSURE		
LINKS /NODES		
Severe injuries/ Fatal accidents/ shpwrecks		
Space consumption of rail link/ noise emission/ derailments		

Inland waterways attributes

List of attributes (most relevant)	PC Navigo	Dutch ministry of transport and NEA	PAWN
Classification Codes			
MOBILE NETWORK ELEMENTS			
TRANSPORTING UNITS			
Vessel categories (based on CEMT 1992 standards), max values:			
TRANSPORT LOADING UNITS			
Containers			
FIXED NETWORK ELEMENTS			
NODES		link type (bridge or lock)	
Lock			
Inland Port			
Level termination point			
Junction			
Intersection			
Transshipment terminals including possible mode shifts			
LINKS			
Class of inl.waterways (CEMT standards):	class of waterway	link type (waterway)	
Type of waterway:			
HIERARCHIES			
NODES			
Hierarchy of transshipment terminals:			
LINKS		category or the waterway	
classification			
Variables and indicators			
GEO-POSITIONS			
Geo-references of nodes and terminals	coordinate of the eastern longitude and of the north latitude		
Name of all types of nodes			
Network section ID			
Link markers			
CONSTRUCTION DATA			
LINKS/NODES			
Costs of construction			
Year of construction (or year of last full renovation)			
Year of last CEMT classification			
Length of section built (in metres for locks)		distance in decametres	distance in decametres
Co-ordinates of section from/to			
Method of financing			

List of attributes (most relevant)	PC Navigo	Dutch ministry of transport and NEA	PAWN
EU contribution if any			
FLOW SUPPORT FACILITIES			
NODES			
Handling of different transportation units			
CAPACITY			
NODES			
Cargo handling capacity		passing time functions	passing time functions
Container handling capacity		for locks and bridges	for locks and bridges
Lock capacity			
Length of chambers			
Depth of chambers			
Quay area, extension of			
Delay time (lock)			
LINKS			
Speed	strenght and direction	speed of the water (for rivers)	speed of the water (for rivers)
Time	of the current	speed of a standard ship	speed of a standard ship
Design inl. Waterway link volume			
Vessel width capacity		capacity of the waterway	
Draught capacity			
Viable for dangerous goods transport			
Varying draught capacity due to changing water levels			
RESTRICTIONS ON USE			
NODES			
Opening hours			
limits to sizes and draught of vessels			
LINKS	permissible and maximum	minimum/maximum loading	maximum loading
Max height for container stacking	passage dimensions of	capacity class of ships	capacity class of ships
Speed restrictions	ships	maximum allowed speed of	maximum allowed speed of
Opening hours	operating hours of	loaded and empty ships	loaded and empty ships
	constructions		maximum draught
FLOW VOLUMES			
NODES			
Goods loaded and/or unloaded splitted into different transportation modes (both weight as value)			
LINK			
Traffic load			

List of attributes (most relevant)	PC Navigo	Dutch ministry of transport and NEA	PAWN
ENVIRONMENTAL PRESSURE Severe injuries/Fatal accidents/Shipwrecks Emissions/ noise			

REFERENCES

The Consortium has produced the following Deliverables and reports:
Del. 1 + 5: *Analysis of existing methods, flows and infrastructure databases*
Del. 2 + 3: *Identifying data on goods and passenger flows*
Del. 6 + 7: *Definition of networks to be reported and selection of network for the database*
Del. 4 + 8: *Data architecture and strategic management. Computer models and data architecture*
Del. 9: *Ensuring compatibility between supply and demand side methodologies*
Del. 10: *Transalpine Pilot Study and demonstrator.*

Report, Workshop 1, May 1998, Copenhagen
Report, Workshop 2, June 1999, Rotterdam

Filename:	Mesudemo FR
Directory:	T:\j\jensemo
Template:	C:\Program Files\Microsoft Office\Templates\Normal.dot
Title:	Final report
Subject:	
Author:	jensemo
Keywords:	
Comments:	
Creation Date:	25/01/02 11:27
Change Number:	2
Last Saved On:	25/01/02 11:29
Last Saved By:	jensemo
Total Editing Time:	1 Minute
Last Printed On:	31/01/02 16:55

As of Last Complete Printing

Number of Pages:	156
Number of Words:	40.894 (approx.)
Number of Characters:	278.085 (approx.)