PUBLIC TRANSPORT
NETWORK TIMETABLE EXCHANGE
(NeTEx)

INTRODUCTION

CEN TC278/WG3/SG9 NeTEx PT001

Fabrizio Arneodo

OCTOBER 2015

TABLE OF CONTENTS

INTRODUCTION

This white paper provides an overview of the CEN Network Exchange standard (NeTEx) for public transport data. NeTEx can to be used for the interoperable exchange of public transport passenger information between systems, supporting a multimodal approach to public transport (including rail, bus, coach and metro systems). It is intended as a European wide standard and mainstream XML technology.

AUDIENCE

NeTEx covers a large and complex domain, and is itself a large standard; this report is intended only to provide a high level introduction for technical managers; for further resources see Further information at the end.

NeTEx IN A NUTSHELL

The NeTEx defines a standard for exchanging public transport passenger information data in XML format. The functional scope of NeTEx is divided into three parts, each covering a functional subset of the CEN Transmodel conceptual model for Public Transport Information, [T1], [T2], [T3].
Part 1 [N1]] describes the fixed **Network** (stops, routes, lines, etc.); Part 2 [N2] is mainly focused on **Timetables** and Part 3 [N3] covers **Fare** data (and is the main subject of this paper). All three parts use the same framework of reusable components, versioning mechanism, validity conditions, support to allow the uniquely identification of data elements in a global context, etc., defined in Part 1. NeTEx also includes container elements called "VERSION FRAMES" to group data into coherent sets for efficient exchange.

NeTEx deliverables comprise (i) a CEN Specification document (in three parts), (ii) a data model in the standard UML modelling language [U1] and (iii) an accompanying XML schema providing a formal electronic description that can be used by data processing software.
Data in NeTEx format is encoded as XML documents that must conform exactly to the schema – standard XML validator tools can check conformance automatically. The schema can also be used to create bindings to different programming languages, automating part of the implementation process for creating software that supports NeTEx formats. Some example XML document encoding different data sets and exchange functions are provided along with the schema.

In effect, documents in NeTEx format are computer files that can be exchanged by a wide variety of protocols (http, ftp, email, portable media, etc). In addition, a SIRI based protocol is specified for use by online web services. The common SIRI framework is used to describe a specific NeTEx/ data service (SIRI-NX) with specialized messages that can be used to request and return messages containing data in NeTEx format, as well as publish/subscribe messages for push distribution. The SIRI-NX responses return a NeTEx XML document that satisfies the request criteria (and also conforms to the NeTEx schema). There is a WSDL binding for this SIRI NeTEx service to make it easy to implement services and service clients as http requests.

A NeTEx service need only implement those elements of relevance to its business objectives – extraneous elements present in the binding can be ignored. Parties using NeTEx for a particular purpose will typically define a "PROFILE" to identify the elements that must be present and the code sets to be used to identify them, for example a French NeTEx profile has been defined that specifies the use of NeTEx for the exhange of NeTEx data.

Whereas TAP/TSI uses optimized flat files that aggregate different fare conditions and prices into a small number of records with dense semantics, NeTEx uses a parameterized approach, with discrete atomic elements that may be combined in many different ways and a ready-made library of almost all known fare conditions. This gives a high level of reuse, and richer semantics, that is, the ability to capture more complex conditions and additional types of fare - but requires a greater effort to understand in the first place. NeTEx uses a uniform design style and coding conventions, which, once grasped, helps to reduce the learning curve.

MOTIVATION

BUSINESS DRIVERS

Modern public transport services rely increasingly on computerized information systems for passenger information; for example for timetables, for real time data and for ticketing. The increased use of online engines and electronic ticket products in particular requires the representation of timetables, products and fares as digitalized data sets. Such data is typically both inherently complex (since the real-world domains it describes are complex) - and subject to a complex workflow. Data is typically assembled from many different stakeholders with different responsibilities (for stops, timetables, real time, fare products, pricing, etc.) and is continually changing at intervals, ranging from the intermittent periodic change of network and timetable data, to the second by second changes of real-time systems.

In this perspective, standardisation seeks to provide effective data models that both capture these complex domains as “REUSABLE COMPONENTS” and to support a workflow that involves continuous integration and validation of data under many different possible configurations of participants.

Well-defined, open interfaces therefore have a crucial role in improving the economic and technical viability of Public Transport Information systems. Using standardised interfaces, systems can be implemented as discrete pluggable modules that can be chosen from a wide variety of suppliers in a competitive market, rather than as monolithic proprietary systems from a single supplier.
Furthermore, individual functional modules can be replaced or evolved, without unexpected breakages of obscurely dependent function. Interfaces also allow the systematic automated testing of each functional module, vital for managing the complexity of increasing large and dynamic systems.

TECHNICAL DRIVERS

Increasing complexity is itself a barrier to the development and uptake of systems, and it is not uncommon to find that organisations develop multiple and sometimes conflicting sub-systems to handle different aspects of their business processes. Complexity also results in difficulty of changing the system, which impairs development of the business. Because PT data sets are complex and shared by so many participants, they are especially hard to change and they thus represent a strategic investment. Therefore it is important to design them for long term use so that they are expressive enough to capture business requirements and flexible enough to evolve to meet to changing business requirements and use.

CEN STANDARDS CONTEXT

CEN (see Figure 1) is Europe's standardisation body. It divides its work into committees covering different aspects of industry and technology. NeTEx, as a transport standard is formally produced by Technical Committee 278, Work Group 3, Sub Group 9. Other TC278 WG3 sub groups handle the related standards Transmodel (SG4) [T1] [T2] [T3], SIRI (SG5 Service interface for real-time information) [S1], and IFOPT (Identification of Fixed Objects in Public Transport [T4]. NeTEx has thus both a concrete standard, and an open consultative process for maintaining that standard.

NeTEx has been developed under the aegis of CEN and is the most recent development stage in over fifteen years work to systemise and harmonize European passenger information data. The work draws on a number of existing national standards applying systematic principles of information architecture to construct flexible models that correctly separate the different concerns of representing and managing data. The keystone is the Transmodel standard, a conceptual model, which names and represents PT information concepts for a wide set of functional areas and can be used to compare and understand different models. Transmodel project outputs have been used both to underpin a number of CEN concrete data standards such as SIRI and IFOPT, and to rationalise national standards to allow for harmonisation and interoperability. Transmodel has been used to develop NeTEx and it is also being updated to include NeTEx additions. Whilst there are a number of standards available for Timetables, NeTEx is the first systematically engineered standard that also covers multimodal Fares.

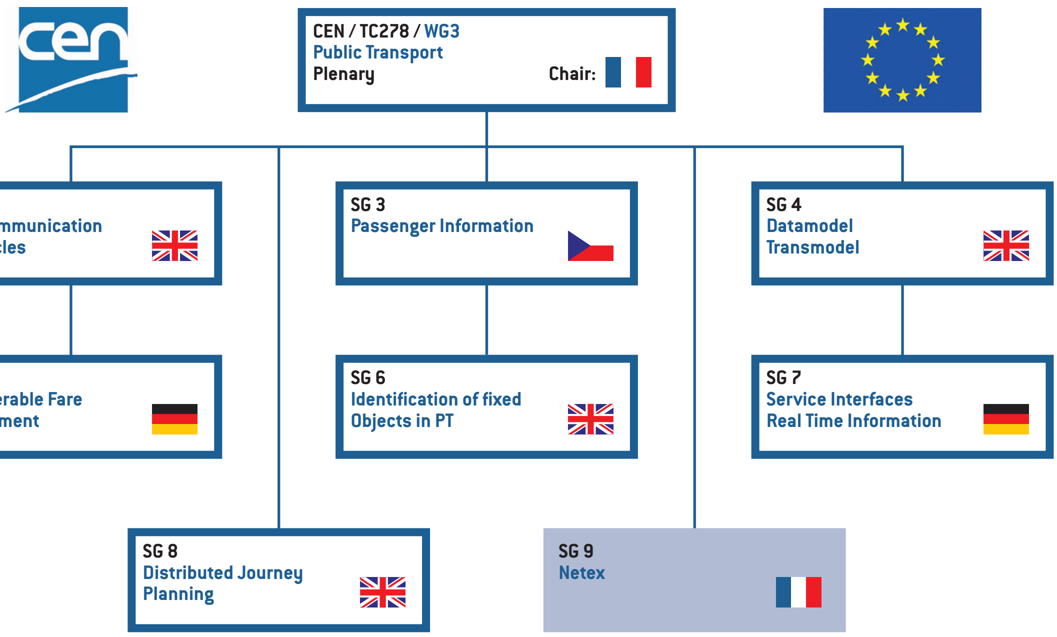

Figure 1 - CEN TC278 WG3 Sub-groups

CEN PROCESS AND PARTICIPANTS

The CEN process provides an organised, transparent method for the governance of standards. It requires a working subgroup to develop a candidate specification, which is then sent to national mirror groups for review and comment, with voting stages for approval and adoption. Work on NeTEx has involved delegates from Austria, France, Germany, Hungary, Italy, Netherlands, Slovenia, Sweden, Switzerland, UK, and the European Rail Authority. Part1 & Part2 were approved in 2013 and Part3 is being finalized at the moment. The involvement of different stakeholders bring to an evolution of the standard as described in the Figure 2.

The development of NeTEx has drawn on existing national and legacy standards such as VDV 452 (DE), BISON (NL) Neptune (FR) and TransXChange (UK), in particular to validate the NeTEx model by establishing mappings with established national standards.
The development of NeTEx also coincided with an interest by the European Rail Authority and other stakeholders in seeking a degree of data interoperability between different modes of public transport such as rail, metro and bus, that is the ability to exchange data about routes, timetables and fares between systems, and also to supply external third party users. To this end a study was undertaken to compare the TAP/TSI B1, B2 and B3 models with the original Transmodel fare model used as the basis for NeTEx Part3; a number of gaps were identified and addressed. A successful informal mapping of the MERITS data for stop and timetable data had already been achieved in Part2.

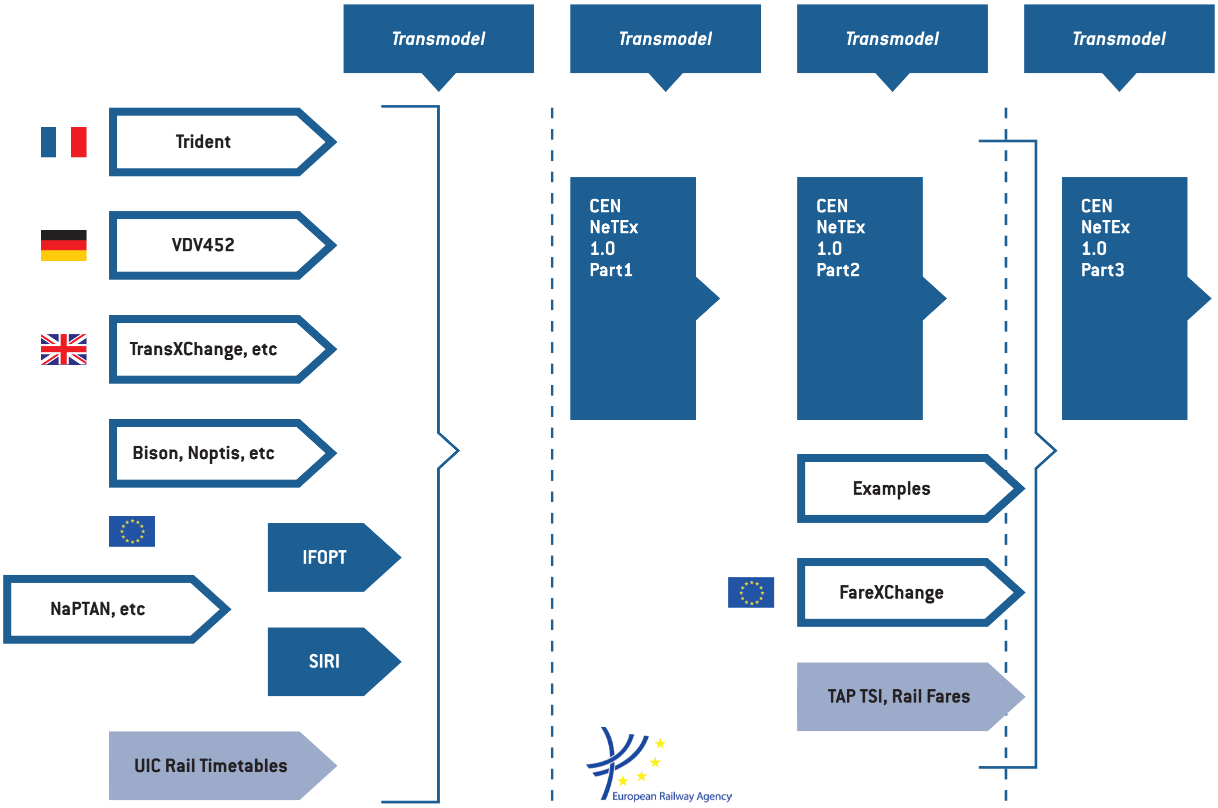

Figure 2 – Evolution of NeTEx Standard

NeTEx RELATION WITH TRANSMODEL

As mentioned, NeTEx is primarily dedicated to data exchange, i.e. an XML message format and a protocol are specified. The content model of the NeTEx message structure is based upon the NeTEx physical data model and is derived directly from Transmodel, the CEN Public Transport Reference Data Model developed at a conceptual level and independent of an implementation in any specific technology (http://transmodel-cen.eu/).

As regards functional domains, NeTEx covers only a subset of Transmodel; the Network Topology, Timing Information, Vehicle Scheduling and Fare Information domains, whereas the full scope of Transmodel is broader, including in addition: Operations Monitoring and Control, Fare Management (sales, validation, control), Management Information and Statistics, Driver Management, Driver Scheduling, Rostering, and Driving Personnel Disposition.

The Transmodel conceptual model is broken down into modular packages with a mostly linear dependency graph between modules. The same organisation of packages is used in the Physical model and XML schema so that there is a direct correspondence between the modules for each functional domain. This makes it straightforward to relate the high level design to the implementation.

THE RELATION OF NeTEx TO GTFS

Google's General Transport Feed Specification (GTFS) [G1] is a widely used format for distributing timetables to third parties. The NeTEx and GTFS formats should be considered as complementary, covering different stages in the data management process, and different workflows: NeTEx is "upstream", GTFS is "downstream". NeTEx differs from GTFS in that it has a much wider scope, and that it is intended for use in multi-sourced back office use cases under which data is generated, refined and integrated (requiring the exchange of additional elements used to construct the timetable), rather than just for provisioning journey planning systems (the prime purpose of GTFS).

GTFS covers stops, lines, and timetabled journey information (Gtf trips) sufficient to answer basic journey planning queries. It supports only a few simple types of fare product. GTFS data identifiers are specific to each data set and require registration of an identifier with Google.

As well has having a much richer fare model, sufficient to cover complex urban and rail products, NeTEx covers many other aspects of Public Transport Information apart from timetables (e.g. network descriptions, interchanges etc,) as well as supporting a richer timetable model for export, one that can include journey patterns, timing patterns, joined journeys, train makeup, connection timings, etc. This makes it able to exchange the component data sets used to build timetables and other operational data sets as well as the resulting timetables themselves. NeTEx includes the additional information needed to provision real-time systems (such as destination displays) and to link to operational systems (such as blocks). It also includes versioning and validity condition mechanisms to support the repeated peer-to peer integration of many data from many different parties.

Because it uses XML, NeTEx is able to package a complete data set as a single coherent document that can be managed and validated.

GTFS uses a traditional flat file format; this is compact and efficient but requires multiple files to describe the different types of element and thus additional rules for naming and packaging the files as a zip. Custom written tools are need to interpret and process the data.

It is possible to generate a full GTFS data set from NeTEx but not vice versa. The NeTEx UML includes a GTFS mapping package which shows how each GTFS element may be populated from the corresponding NeTEx element.

ACKNOWLEDGEMENT

This White Paper has been produced thanks to the support of Fabrizio Arneodo (5T srl), Kasia Boureé (KBIC), Christophe Duquesne (Aurige), Nick Knowles (Steam Intellect), Andrej Tibaut (University of Maribor), Jan Tijmensen (InTraffic , Connekt).

FURTHER READING

THE NeTEx STANDARD

[N1] NeTEx-Part 1: *Public Transport Network Topology exchange format, CEN/TS 16614-1:2014*
[N2] NeTEx-Part 2: *Public Transport Scheduled Timetables exchange format, CEN/TS 16614-2:2014*
[N3] NeTEx-Part 3: *Fare Information exchange format, CEN/TS 16614-3:2014*

OTHER NeTEx WHITE PAPERS

[W1] NeTEx *Getting Started* - White Paper
[W2] NeTEx *Design Methodology* - White Paper
[W3] NeTEx *Framework* - White Paper
[W4] NeTEx *Reusable Components* - White Paper
[W5] NeTEx *Network* - White Paper
[W6] NeTEx *Flexible Networks and Multimodality* - White Paper
[W7] NeTEx *Accessibility* - White Paper
[W8] NeTEx *Timetables* - White Paper
[W9] NeTEx *Fares* - White Paper

OTHER REFERENCES

[T1] *Public Transport Reference Data Model - Part 1: Common Concepts (Transmodel)*, EN12896-1
[T2] *Public Transport Reference Data Model - Part 2: Public Transport Network (Transmodel)*, EN12896-2
[T3] *Public Transport Reference Data Model - Part 3: Timing Information and Vehicle Scheduling (Transmodel)*, EN12896-3
[T4] *Identification of Fixed Objects for Public Transport*, EN28701
[S1] Service Interface for Real Time Information (SIRI) prCEN/EN EN 00278181, prCEN/EN EN 15531
[U1] *Unified Modeling Language* **www.omg.org/spec/UML/2.5/**
[G1] *General Transport Feed Specification* https://developers.google.com/transit/gtfs

The World Geodetic System (WGS) is a standard for use in cartography, geodesy, and navigation including by GPS. It comprises a standard coordinate system for the Earth, a standard spheroidal reference surface (the datum or reference ellipsoid) for raw altitude data, and a gravitational equipotential surface (the geoid) that defines the nominal sea level. The latest revision is WGS 84 (aka WGS 1984, EPSG:4326).

FURTHER INFORMATION

NeTEx website: **www. netex-cen.eu**
Enquiries / email: **info@netex-cen.eu**