Towards Knowledge Modeling for Sustainable Transport

Paloma Cáceres(✉), Carlos E. Cuesta, José María Cavero, Belén Vela, and Almudena Sierra-Alonso

VorTIC3 Research Group, Rey Juan Carlos University, C/Tulipán, s/n, 28933 Móstoles, Madrid, Spain
{paloma.caceres,carlos.cuesta,josemaria.cavero, belen.vela,almudena.sierra}@urjc.es

Abstract. The paradigm shift from the current energy consumption model towards a sustainable model requires to develop new behaviors and strategies. This is particularly relevant in domains like the public transport. Many providers are currently offering services to assist passengers to plan their routes. However, these approaches are often restricted to some specific area or transport medium. We suggest using a Linked Data perspective, which makes simpler to combine data from different sources, as well as extending and managing them. Moreover, it makes possible to enrich the basic model to the extent of developing a knowledge model, able to use semantic techniques to unfold even better strategies. In this paper we present a proposal in the transport domain, which refines a basic model into a Transmodel specification and later adds more information according to the IFOPT model. This defines a knowledge model, which can be used to develop sustainable transport strategies.

Keywords: Sustainable development · Knowledge management · Semantic web · Transmodel · IFOPT · RDF

1 Introduction

Sustainable development makes reference to a mode of human development in which the use of resources aims to meet human needs while ensuring the sustainability of natural systems and the environment, so that these needs can be met not only in the present, but also for generations to come [1].

In this context, transport represents a significant proportion of global energy consumption. Sustainable mobility essentially emerges from the disadvantages caused by the current transport model, in which the central element is the private car. These include pollution and its effects on health and the environment, inefficient use of resources, traffic congestion, etc. These disadvantages have triggered different efforts to search for alternatives, trying to overcome the limitations of this model. In many cases, energy consumption can be significantly

S. Counsell and M. Núñez (Eds.): SEFM 2013 Collocated Workshops, LNCS 8368, pp. 271–287, 2014.
DOI: 10.1007/978-3-319-05032-4_20,

reduced by an efficient use of transport media. In this way, there are many reasons to promote the use of public transport as environmental and economic arguments [2].

It is important, not only choosing the right path within the transport network, but also being to share a vehicle (including private vehicles) with others. There are several initiatives assisting people in sharing transport, the most popular being *carpooling*: more than one person sharing a car. However, no existing solution combines private transport sharing with the use of public transport, which would make it more flexible.

Having this in mind, we aim to apply information technologies to the task of improving citizen mobility – considering the case of citizens, both in usual or specific trips, and trying to optimize their intended routes by using any available means of transport: both public transport and the rational sharing of private transport. We have designed an IT platform, called CoMobility [3], to assist in intermodal transport sharing, integrating the use of carpooling with public transport, as well as other private transport media. To be accessible anytime and anywhere, we have taken into account mobile computing. Mobile systems are not only useful in their ability to provide pervasive access to computing systems (i.e. enabling computer access anywhere), but they also provide the inputs of a mobile and dynamic environment into a computational system; it is now possible to perform computations that, until recently, were simply impossible. A classic example is geolocation: it is now easy to provide the physical location of a user, and to use these data for a variety of purposes.

It is provided on the Internet "as a service", where both public transport information and data provided by users themselves are stored and accessed "in the cloud". The cloud approach is necessary as *scalability* is one of the most important requirements of this kind of wide-range service architecture. The platform also needs to access *a great amount of data*, which is also stored in the cloud – both the private data of carpoolers, and the public data accessed in a linked open data approach. Users are able to access their information in several formats, particularly in mobile devices (currently, Android devices) and web applications. Through these devices, they are able to plan their paths in the transport network, moving from a shared car to the underground, and from there to a bus line; and at the same time receiving an estimation of the saving of both money and energy. For this purpose, our CoMobility platform has "customized" analytics on savings and energy consumption, to make individuals aware of the benefits of this new way of travelling. These data are obtained from energy-aware institutions.

This work focuses on how to incorporate new knowledge into open data about transport, provided by its original sources. The format of public data within the open data initiatives prevents non-experts from using them directly, and thus it requires additional semantics, as provided by *"Linked Open Data"* [4,5]. In this way, our proposal in this paper focuses on modeling the knowledge of transport public data for sustainable transport. To do that, it is necessary: first, analyzing the original data formats and identifying the data semantics; second, matching these data with the vocabulary of the transport metamodels Transmodel [6] and IFOPT [7], and adding the relevant information; and finally, representing them as linked open data.

The paper is structured as follows: in Sect. 2, we briefly introduce the resource description framework (RDF). Section 3 outlines the public transport data providers (in this case, the public bus company of Madrid) and the transport specification standards Transmodel and IFOPT. Section 4 describes our specific proposal about a knowledge management architecture for Public Transport; in Sect. 5 we provide an example based on our proposal; and finally, the main conclusions are shown in Sect. 6.

2 A Brief Introduction to RDF

The Resource Description Framework (RDF) [8] provides an extremely simple data model in which entities (also called resources) are described in the form of triples (subject, predicate, object). For instance, consider a meteorology system consisting on several sensors, and a given sensor identified as `Sensor_4UT`; its description could comprise the following triples:

```
(Sensor_4UT, rdf:type, om-owl:System)
(Sensor_4UT, om-owl:parameter, weather:_AirTemperature)
(Sensor_4UT, om-owl:parameter, weather:_RelativeHumidity)
(Sensor_4UT, om-owl:parameter, weather:_WindDirection)
(Sensor_4UT, om-owl:parameter, weather:_WindSpeed)
```

The first triple states that `Sensor_4UT` is a particular class of system (hence the object `om-owl:System`), and the remaining four triples say that `Sensor_4UT` reports measurements about air temperature, relative humidity and wind direction and speed.

An RDF dataset can be seen as a graph of knowledge in which entities and values are linked via labeled edges. These labels (the predicates in the triples) own the semantics of the relation, hence it is highly recommendable to use standard vocabularies or to formalize new ones as needed. Figure 1 represents the previous RDF excerpt as a labeled graph in which the nodes (and edges between them) depict the mentioned triples, modeling a weather observation about the wind speed.

RDF has been gaining momentum since its inception thanks to its adoption in diverse fields, such as bioinformatics, social networks, or geographical data. The Linked Open Data project plays a crucial role in the RDF evolution [4]. It leverages the Web infrastructure to encourage the publication of such semantic data [7], providing global identity to resources using HTTP URIs. Moreover, integration between data sources is done at the most basic level of triples, that is, to connect two data sources can be as easy as making connection between those resources.

This philosophy pushes the traditional document-centric perspective of the Web to a data-centric view, emerging a huge interconnected cloud of data-to-data hyperlinks: the *Web of Data*. Latest statistics pointed out that more than

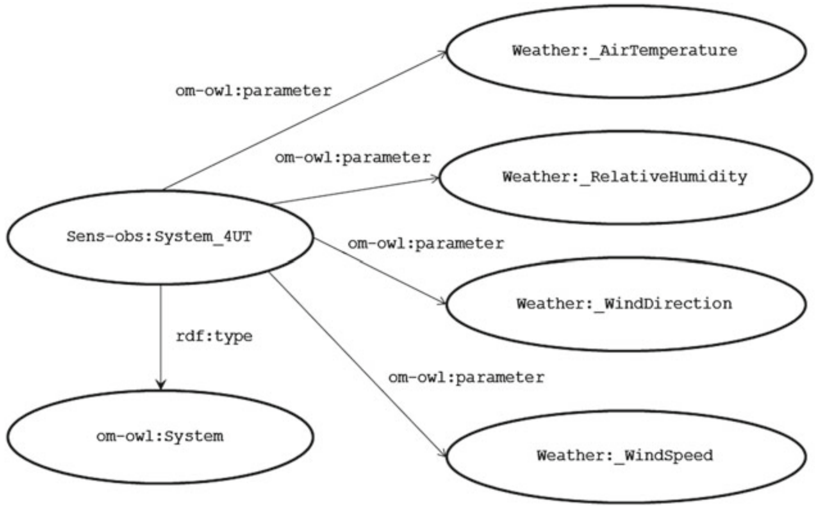

Fig. 1. Example of a RDF graph modeling weather data

31 billion triples were published and more than 500 million links established cross-relations between datasets. Although each piece of information could be particularly small (the so-called Big Data's *long tail*), the integration within a subpart of this Web of Data can also be seen as an example of Big Semantic Data.

It is worth noting that RFID labels, Web processes (crawlers, search engines, recommender systems), smartphones and sensors are potential sources of RDF data, such as in our previous use cases. These are the main players in the so-called *Internet of Things*, in which the Linked Data philosophy can be applied naturally by simply assigning URIs to the real-world things producing RDF data about them via Web. As a result, the activity of all involved devices is recorded and linked between them, enabling large projects (such as the emergent notion of smart-cities) to be successfully implemented.

3 Using Knowledge Modeling for Sustainable Transport

This work is focused on modeling public transport data in Linked Open Data (LOD). In this way, our work aims to study the source of the public transport data and the existing standard models related to public transport, to analyze them and to enrich this information adding semantic knowledge of the public transport data. The final data will be published in RDF.

3.1 A Modeling Approach to Public Transport

In this paper, we work with the public transport data of the public bus network in Madrid as open data source. EMT Madrid [9] is the public bus company of

Madrid, and follows an open data initiative with regard to its information. It provides the specification of the geographical services public platform, which includes these three parts: a customer-oriented SOA architecture [10], a public information services infrastructure architecture –described in a detailed way–, and the description of their information services. These services can be accessed by different kinds of devices.

EMT Madrid allows accessing its platform by means of the explicit authentication of users. This method is implemented in each Web Service. Moreover, it provides a specific set of output information about the bus line and the bus stop, as a set of fixed schemas into the services structure:

- With regard to the bus line, EMT Madrid offers a piece of basic information which can be shown on any device screen: Number of bus line, first/last bus stop and last/first bus stop (depending on the direction followed by the bus), type of the day in which the bus operates (to indicate working day, holiday and Saturday), direction, start and finish time of the bus line (in that direction), maximum and minimum frequency of the bus line (in that direction).
- With regard to the bus stop, EMT Madrid also offers a piece of basic information which can be shown on any screen, including those from mobile devices: code, name, postal address, geographical coordinates.

We have studied the standard transport models Transmodel [6] and IFOPT [7], to match the existing open data about the bus public network of EMT Madrid with them. Transmodel is an European Reference Data Model for Public Transport Information which provides a model of public transport concepts and data structures that can be useful to build information systems related to the different types of public transport. It includes information about real time data, journey planning, timetables, operational management, routes, etc. The present version (V5.0) uses an Entity-Relationship modeling approach and covers the following domains:

- Tactical Planning,
- Personnel Disposition,
- Operations Monitoring and Control,
- Passenger Information,
- Fare Collection and
- Management Information/Statistics

Transmodel establishes a consistent terminology for describing public transport concepts, providing definitive equivalents for use in the National Languages of each participant nation. Where public transport (PT) related words in vernacular use may span a number of different concepts and lead to differences of interpretation, it establishes a more precise technical terminology for unambiguous use by PT information system developers. For example the terms "trip", "journey", "service", are overlapping concepts that in Transmodel are used only in some more specific usages.

Sometimes, we need more descriptive information about the objects related to public access to Public Transport than the offered by Transmodel. Let's see an example: a Transmodel Connection Link represents the possibility of interchange between two Scheduled Stop Points used by different journeys, *without necessarily having a precise indication of place.* In contrast, a Path Link (from the IFOPT metamodel) represents a different information layer: a Stop Path Link and an Access Path Link represent the possibility of navigation between specific located nodes of a Stop Place. A Transmodel Access Link is the physical (spatial) possibility for a passenger to access or leave the PT system: the walking movement of a passenger from a place (origin of the trip) to a stop point (the origin of the PT trip); the walking movement of a passenger from a stop point (the destination of the PT trip) to a place (destination of the trip.

For this reason, we have studied the IFOPT metamodel. It defines a model and identification principles for the main fixed objects related to public access to Public Transport (e.g. stop points, stop areas, stations, connection links, entrances, etc.). The IFOPT Standard builds on the TransModel Standard to define four related submodels:

- *Stop Place Model*: Describes the detailed structure of a Stop Place (that is station, airport, etc.) including physical points of access to vehicles and the paths between the points, including mobility hazards.
- *Point of Interest Model*: Describes the structure of a Point of Interest including physical points of access, i.e. Entrances.
- *Gazetteer Topographical Model*: Provides a topographical representation of the settlements (cities, towns, villages etc.) between which people travel. It is used to associate Stop and Station elements with the appropriate topographic names and concepts to support the functions of journey planning, stop finding, etc.
- *Administrative Model*: Provides an organizational model for assigning responsibility to create and maintain data as a collaborative process involving distributed stakeholders. Includes namespace management to manage the decentralised issuing of unique identifiers.

Our study focuses on the Stop Places model of IFOPT.

3.2 IFOPT: Modeling Stop Places

What type of information can be better modeled using IFOPT instead of using Transmodel? In the previous subsection, we introduced an example of related information between both metamodels and their differences. Now, we will show the differences graphically, explaining them in depth.

In Fig. 2, we show a route from a starting point (numbered as 1) to a destination (numbered as 6). This route is composed by three pedestrian subroutes (1-2, 3-4 and 5-6) and two PT subroutes (2-3 and 4-5). The pedestrian subroutes 1-2 and 5-6 are identified as Access Links (paths to access to PT) in Transmodel. The pedestrian subroute 3-4 is a Transmodel Connection Link (to connect different PT routes).

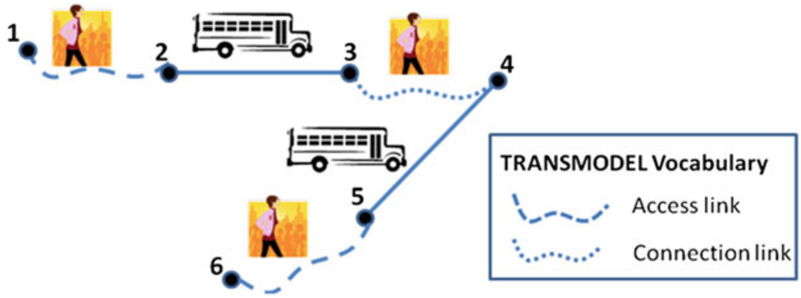

Fig. 2. Transmodel access and connection links

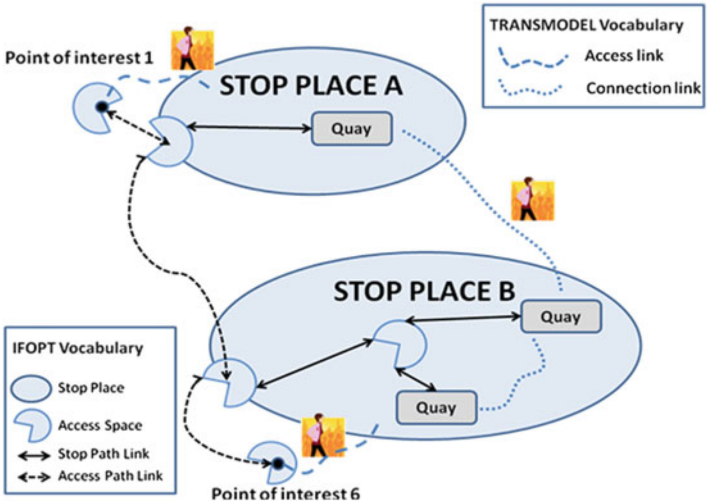

Fig. 3. Differences between transmodel and IFOPT elements

IFOPT is more specific describing fixed objects. It models Stop Places, Access Path Links and Stop Path Links. Stop Places are related with a place where users can get on public transport. A Stop Place includes an Access Space which provides the entrance to the place. The place could also be a Point of Interest, for example. Stop Path Links represent the path between stops of public transport, within a Stop Place. Access Path Links represent the path to access to an Access Space from another Access Space.

Figure 3 shows Points of Interest (numbered as 1 and 6), Stop Places (labelled as A and B), Access Path Links and Stop Path Links, which are represented as arrows. Each Stop Place and Point of Interest has an Access Place which represents the entrance to these places. We want to emphasize that the Stop Place B also includes an Access Space which could represent the space where users could change the journey or the public transport. We can also see in the figure that a Transmodel Access Link is less descriptive than an Access Path Link, which represents the path through an Access Place. IFOPT's Stop Path LINKs represent specific paths from/to an Access Space to/from a public transport stop, including transfers within Stop Places.

4 A Knowledge Management Architecture for Public Transport

4.1 Context: The CoMobility Project

The CoMobility Project [3] defines a multimodal architecture based on linked open data for a sustainable mobility. Its main goals are improving the citizen mobility, optimizing their trips combining both public transport and sharing private transport (i.e. car sharing or carpooling), providing accessible trips when necessary and saving energy and reducing the pollution (Fig. reff.comobility).

We have developing a systematic approach to (i) accessing open, integrated and semantically annotated transportation data and street maps, (ii) combining them with private data, and (iii) supplying mechanisms to allow the actors to share and search these data. Therefore, its conceptual architecture provides the means to perform the following tasks: First, the platform can identify, select, extract and integrate data from different and heterogeneous sources, stemming from the transportation, geographical and energy domains. Second, data from public institutions is obtained automatically in the form of open data. Third, these data are annotated as linked data, and a set of heuristics generate links between data items from different sources without human intervention. Fourth, these data are integrated with private data provided by users themselves. And

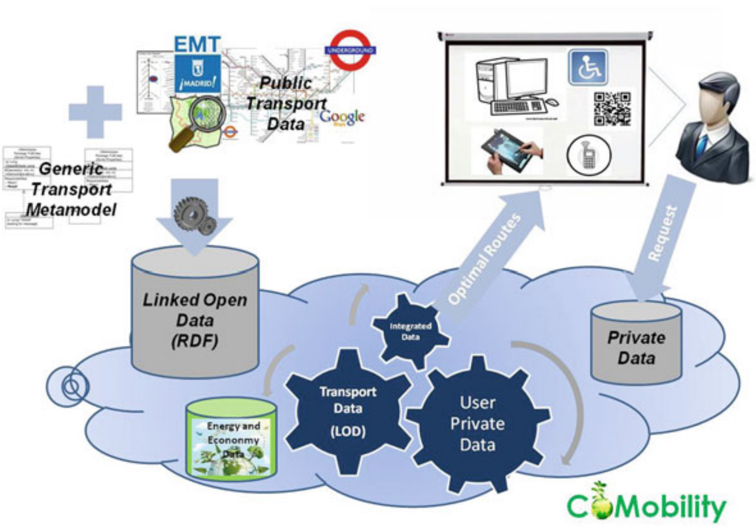

Fig. 4. The CoMobility Project

finally, CoMobility provides intuitive and customized data analytics and visualization, allowing individuals to become aware of the environmental impact of their transport choices. Next figure shows a general idea about the project.

Project CoMobility is supported by the Spanish Economy and Competitiveness Ministry and some companies have expressed their interest in their results. The most representative are: the public bus company from Madrid (EMT Madrid), the Public Regional Consortium of Transports of Madrid (CRTM) [11], the Chair of EcoTransport, Technology and Mobility of the Rey Juan Carlos University [12] and the Spanish National Society of Blind People (ONCE) [13].

4.2 Knowledge Management Architecture

In this paper we briefly describe the knowledge management architecture that supports the bus public transport data from EMT Madrid as linked data in RDF. To carry out this process we need: first, to provide the required mechanism to parse the XML [14] information from the source and to obtain the necessary data about the network transport (i.e. bus lines and bus stops); second, to mach these data with the IFOPT model; and finally, to publish them as RDF resources. Next subsections show the requirements of the architecture which supports this process.

4.3 Data Architecture for the Original Data

EMT Madrid provides an open data platform of its urban bus network, supported by a technological infrastructure which exports them as a Service-Oriented Architecture (SOA) to any requesting consumer (any consumer that requires it). In this way, the data can be accessed using Web services, which answer to the request by returning some specific XML files. EMT Madrid offers a set of open web services which clients can invoke according to their information needs: bus routes, pedestrian routes, bus stops, transfers, etc. The figure below shows the concrete architecture that support this kind of information interchange (Fig. reff.emt).

Each web service returns a different XML file which contains the information about the requested data by the client. It is then necessary to parse and translate the XML information into a readable information to the client. Figure 6 shows an XML example of a bus line. To preserve the confidentiality, we have hidden some data from the original information, or changed the specific formatting, in the following examples.

The code shows a generic description about the bus line 174 and its headers. First, the labels `<BusLine>` ... `</BusLine>` indicate that the information is about a bus line. The bus line number is 174: `<IdBusLine>174</IdBusLine>` and the first and last stops are Plaza de Castilla and Sanchinarro Este: that is, `<BusLineHeadA> PLAZA DE CASTILLA </BusLineHeadA>`, and then `<BusLineHeadB> SANCHINARRO ESTE </BusLineHeadB>`, respectively.

Figure 7 shows an XML example of a bus stop. The code shows the bus line number to which the bus stop belongs: `<IdBusLine>174</IdBusLine>`. Then, the bus stop number, <Node>5611</Node>; the distance from the first (initial) stop, <Distance> 456 </Distance>; and the distance from the previous stop <DistancePrev> 147 </DistancePrev>; and then the bus stop name, <BusStopName> INTERCAMBIADOR PZA. DE CASTILLA </BusStopName>. Finally, it also specifies the geographical coordinates, i.e. latitude <GeoCoorX>40,4695235553361 </GeoCoorX> and longitude <GeoCoorY>-3,68778542580241</GeoCoorY>.

4.4 Superposing Data Architectures

We need to identify which data from the source are corresponded with the structured of IFOPT. As mentioned before, IFOPT defines a model and identification principles for the main fixed objects related to public access to Public Transport (e.g. stop points, stop areas, stations, connection links, entrances, etc.). In this paper, we only work with the Stop Place Model, because it describes the detailed structure of a Stop Place (that is station, airport, etc.) including physical points of access to vehicles and the paths between the points, including mobility hazards.

As we will see in the next section (by developing an example), we actually need not to define a set of different models to later combine them – instead of that, we have information (models) from several sources which can be easily combined using specific *join points*. We begin with a simple model (i.e. bus

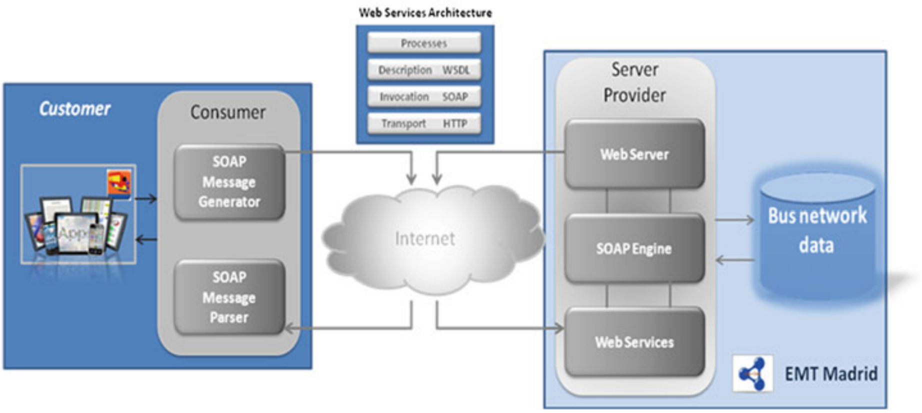

Fig. 5. SOA-based EMT infrastructure

```
<BusLine>
  <IdBusLine>174</IdBusLine>
  <BusLineHeadA>PLAZA DE CASTILLA</BusLineHeadA>
  <BusLineHeadB>SANCHINARRO ESTE</BusLineHeadB>
</BusLine>
```

Fig. 6. Example of a XML-based bus line description

```
<BusStop>
  <IdBusLine>174</IdLine>
  <Order>1</Order>
  <Node>5611</Node>
  <Distance>456</Distance>
  <DistancePrev>147</DistancePrev>
  <BusStopName>INTERCAMBIADOR PZA.DE CASTILLA</BusStopName>
  <GeoCoorX>40,4695235553361</GeoCoorX>
  <GeoCoorY>-3,68778542580241</GeoCoorY>
</BusStop>
```

Fig. 7. Example of a XML-based bus stop description

lines), where the fixed points are described as simple bus stops; and we enrich this model, step by step, adding the information provided by the different model, until these simple stops are transformed into complex structured Stop Places.

5 Exporting the Transport Model as Linked Data

As already noted, we simply enrich the original model. The process can be described as the stepwise refinement of the original dataset provided by our original sources (EMT Madrid), simply provided as or translated to Linked Data (RDF) format, and then enriched with additional semantic information – in the first stage, identifying relevant sections according to the key concepts in Transmodel; then in a second stage, attaching additional information as described by the IFOPT model.

5.1 Representing the Original Information

First, consider the original source structure as provided by EMT Madrid, described in RDF terms. As already noted, as EMT Madrid is a bus service, it provides information in terms of bus lines, bus stops, etc. This can be simply described as Linked Data, essentially by represented the provided information as a graph, and thus reordering some elements in the structure as required.

The most important notion in the Bus model is still that of *route*, i.e. the path which must be followed by a passenger to reach his destination from some starting point. However, for the sake of simplicity, in the remainder of this example we will focus in a small part of the model – specifically, the one which refers to bus stops (and hence to bus lines, and to related pedestrian walks).

The first notion that we need to define is that of a *bus line*, which describes the first intuitive concept as provided by EMT Madrid. In RDF terms, a bus line is described as a resource in a certain URI, with a number of triples defining a number of attributes (line identifier, origin name, destination name), and also a number of relationships or connections (prominently, to the bus stops locating the start and the end of the line). Figure 8 depicts this structure in the usual graphical form, for EMT Madrid line 174.

Of course, both the predicates defining attributes (`att:`x) and those defining connections (`conn:`x) are already defined in RDF terms. Both abbreviations are defined as alias for longer URIs – for instance, the prefix `att` in the Figure is the short form for `http://vortic3.com/rdf/attributes#`.

Connection definitions show that bus lines need to refer to bus stops to locate their heads – and even more, any stops within the line. Therefore, bus stops define the obvious connection between lines. Indeed, bus stops can be considered as a key notion in the Bus model, and they provide a starting point to traverse the whole graph of bus lines – i.e. the triplestore.

In RDF terms a bus stop is again described as a resource, with certain attributes (name of the stop, e.g. the direction; spatial coordinates, etc.), and also a number of relationships (prominently, the crossing with bus lines, i.e. the set of lines which stop at this bus stop). This is depicted in the first part of Fig. 9, which shows part of the dataset describing the stop at "Intercambiador Plaza de Castilla".

Of course a stop is able to simultaneously participate in several bus lines – this is the case of the aforementioned "Intercambiador" (i.e. a transport hub). To be able to capture this, the auxiliary resource known as *node* (or crossing) is defined. A node describes the role of a certain bus stop in a certain line. For instance, the second half of Fig. 9 describes the node `1000A4`, which represents the role of the bus stop at "Intercambiador Plaza de Castilla" in the already mentioned bus line 174 (see Fig. 8). Hence the attributes of the node describe the name of the stop, its position in the bus line (it is the first stop), its distance to the previous stop, etc. The node has also relationships, prominently the connection linking it to the corresponding bus line. As depicted in the lower part of the Figure, bus lines have separate definitions when they are considered in the opposite direction

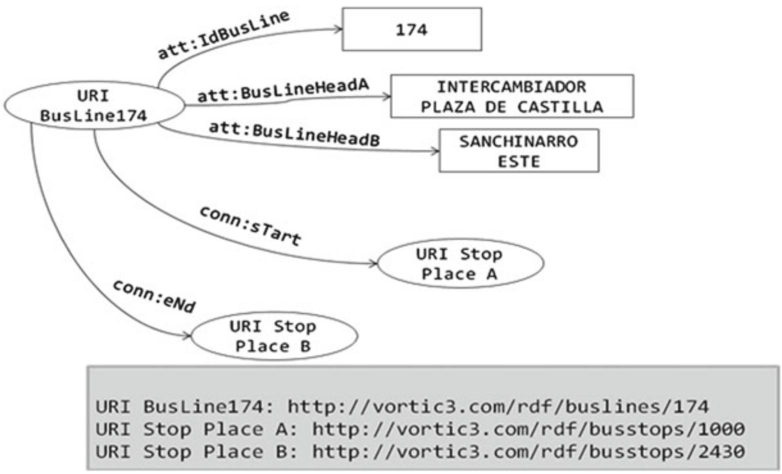

Fig. 8. RDF example for the bus line 174: bus line

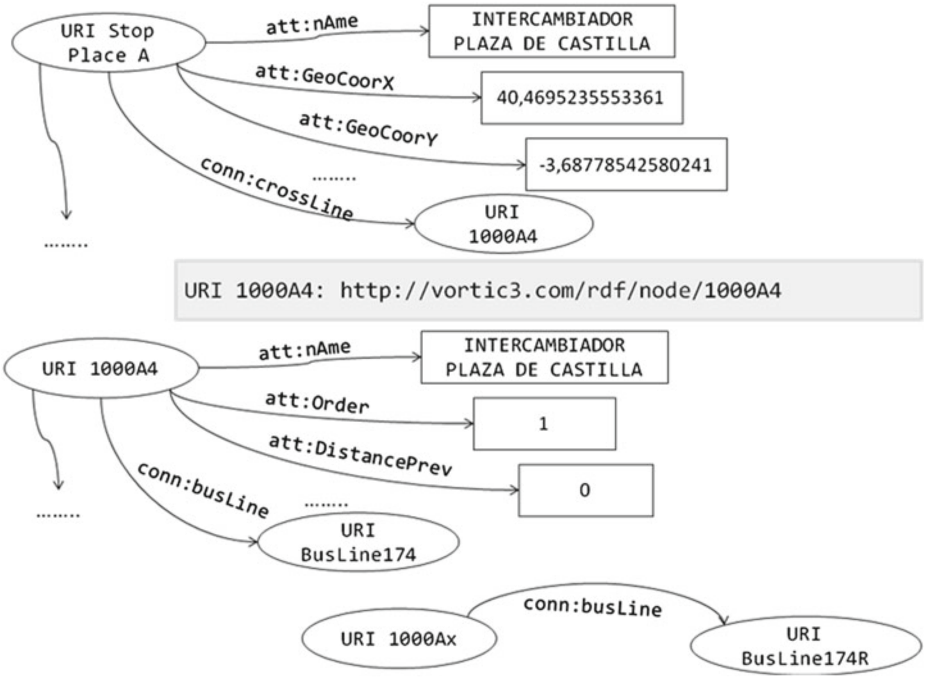

Fig. 9. RDF example for the bus line 174: bus stops and nodes

– and hence, if the same stop participates in this reverse line, the connection is described as a different node (that is, node `1000Ax` in the Figure).

5.2 Exploring Connections in the Network

As already noted, we are particularly interested in the concept of *transfer*. Within the Bus model (and therefore in the bus lines domain) a transfer refers just to a change of bus line – the passenger steps off a certain bus line, and gets into another one. In this model, the number of transfers equals the number of changes – i.e. the number of bus lines in the route equals the number of transfers plus one.

In the initial conception, a transfer can happen when a certain stop crosses two lines, i.e. when it has at least two nodes. That is the simplest transfer: the passenger just leaves the first line, and in the same stop, he takes a different bus line. However, things are often more complicated – the passenger is able to walk to a different stop in the nearby, to do a different transfer as described in Figs. 2, 3. This notion of a *pedestrian walk* implies a different connection between two bus stops, and can be also modeled as a RDF resource, which starts in (i.e. connects to) one stop, covers a distance and (optionally) an average duration (attributes) and ends in (again, connects to) another stop.

So far we have used information from the Bus model as provided (and translated) by EMT Madrid. However, once we are out of the bus for a pedestrian walk, we can generalize the situation: if the passenger is able to walk or to take the bus, he can also use the subway – i.e. we might consider not only transfers between bus lines, but also movements between different PT providers.

5.3 Adding Transmodel Concepts

To generalize this, we just need to get beyond the Bus model, to go to a model when any kind of PT can be used; for this purpose, we are able to use the Transmodel definition to refer to any kind of transport, as already indicated in Sect. 3.1. Therefore the existing bus definitions are substituted by their equivalent (and more general) version in Transmodel. Instead of bus stops we have Stop Points, which could refer to any sort of stop, including Metro or train stations as well as bus stops. The definition of a pedestrian walk in Transmodel terms is summarized as described in Fig. 2. We consider all stop points from a certain origin to a certain destination, and pedestrian walks are captured either as Access Links (i.e. the way we access a certain stop point) or as Connection Links (i.e. a non-fixed path between two stop points). The notion of transfer corresponds to an *interchange* in Transmodel. In summary, the original data in the Bus model fulfills the requirements from a Transmodel specification, hence it can be immediately translated to these terms – and this means that it can be safely extended to include information about other PT media.

Therefore, the original structure is easily translated to a Transmodel specification, including all its data without any modification. Now these data can be easily extended to include e.g. the stop point which describes a Metro station, and hence we can use the connection link between this station and any other stop point to describe how to move from the bus to the subway – or vice versa. The notion of transfer between buses is now generalized to any kind of connection between PT media.

Note that the model is not required to be complete: we can have the full definition of bus lines and stops as provided by EMT Madrid, but we can add only a few Metro stations at the beginning – the connection links would be used wherever they are present, but the semantic model is perfectly able to work in the presence of partial information. That is, we need not to have the full definition of the Metro network to describe a partial combination.

5.4 Introducing the IFOPT Basic Model

However, this is not enough: as already indicated (Sect. 3.1, up to the comparison in Fig. 3), there are many details and attributes, related to different PT media which cannot be adequately described as Transmodel stop points. The Transmodel definition does not provide enough information to be able to handle the most complex reasoning related to simple bus transfers – even more for generic interchanges. Therefore we need to extend this information – hence the need for refinement, and the use of the IFOPT model, the standard which complements Transmodel providing information about fixed points. In particular, the notion of *stop point* is generalized to the concept of Stop Place, as already noted before (note also that even in Fig. 9, we have already modeled the bus stop actually as a stop place).

As already indicated (see again the comparison in Fig. 3) the notion of stop place is more general than a stop point, as it can also include generic Points of Interest. The concept of *access link* is refined into a more structured Access Space, and *connection links* are considered either "internal" (Stop Path Links) or "external" (Access Path Links). Therefore, we can still represent, using the same structure, a bus transfer – i.e. two stop places and the access path link between them; but we are also able to describe how the passenger moves into a communication hub. For instance, the same stop place already described in Fig. 9 ("Intercambiador Plaza de Castilla") includes several bus stops (such as the connection to line 174, described as a *node* in the same Figure), but also a Metro station. This is part of the *same* stop place, but refers to a different *quay*, which would have its own access spaces. Therefore, the transfer between bus 174 and Metro line 10 at Plaza de Castilla can described here as a *stop path link*, as the passenger needs not to go out of the stop place to change the PT media.

5.5 Using Linked Data Information

Obvioulsy, all these notions are described as RDF resources. Bus stops are defined as stop places, as they already were. Metro lines are described similarly to bus lines, according to Transmodel definitions. Stop places are refined into *quays* (generalizing nodes); and instead of having "bus transfers" (which were not described in RDF terms), we actually have *access path links* between access spaces, and *stop path links* between the quays in the same stop place. All these elements are directly modeled as RDF resources with their own attributes, and these links take the form of connections between these resources.

Using the rich semantic model as provided by IFOPT, we are able to model all kinds of situations within this context. Specifically, we are particularly interested in aspects related to sustainable transport or to accessibility of disabled people. For instance, with regard to the first, IFOPT provides the concept of *CheckPoint* as a component of the stop place; attributes related to this checkpoint include estimates about the duration of the delays it might cause (*CheckPointDelay*), and the possibility of *Congestion*. With regard to the second, it is even more apparent, as according to IFOPT, any stop place element can include a validation related to its potential *AccesibilityLimitation*; which is described in turn as a resource with such attributes as *WheelchairAccess* or *VisualSignsAvailable*. Therefore, now we are not just able to tell if there is an access path link from stop place A to stop place B; but also if this path can be safely used by blind or disabled people.

6 Conclusions

This paper exposes two main arguments. First, that it is both simple and convenient to either design or translate the information from different sources –in our case, from the transport domain– to be managed as Linked Data. The generality and flexibility of the RDF format makes possible to express data from almost any source, and to easily translate to it information from varied origins – in the presented example, our data was originally provided as XML files and private services; and in both cases it is simple to encapsulate it as RDF triples. The generic nature of this format makes easy to integrate the information from many sources.

The second thesis is that it is possible to elaborate the information step by step, in a refinement process, starting with the most basic ("raw") data and to enrich these data adding information from semantically rich models – in our context, first the Transmodel specification for transport media, and then the IFOPT model to describe fixed points. The open nature of the Linked Data approach makes possible to perform this stepwise refinement in a seamless way, adding RDF triples without the need to modify the existing ones. The result is a semantically rich model, describing all available knowledge about PT media and the related places, which allows for a sophisticated form of reasoning. The ultimate purpose of our system –to find optimal routes in the PT network– is fulfilled to the extent of being able to optimize these routes considering not only traversability, but also sustainability and accessibility.

Acknowledgements. This work has been supported by the project CoMobility (TIN2012-31104), funded by the Spanish Ministry of Economy and Competitiveness; and it has also been supported by the Chair of Ecotransport, Technology and Mobility (http://www.catedraetm.es/) at Rey Juan Carlos University.

References

1. Sustainable development - Wikipedia. http://en.wikipedia.org/wiki/Sustainable_development (2013). Accessed 21 Nov 2013
2. Essortment Home Page. Advantages of Using Public Transportation. http://www.essortment.com/advantages-using-public-transportation-19689.html (2012). Accessed 22 June 2012
3. Cuesta, C.E., Cáceres, P., Vela, B., Cavero, J.M.: CoMobility: a mobile platform for transport sharing. ERCIM News 93 (2013)
4. Linked Open Data. http://linkeddata.org/
5. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. Morgan & Claypool, Seattle (2011)
6. Transmodel: European Reference Data Model for Public Transport. http://www.transmodel.org/index.html (2001)
7. Identification of Fixed Objects in Public Transport (IFOPT). http://www.dft.gov.uk/naptan/ifopt/ (2013)
8. Klyne, G., Carroll, J.: Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation. http://www.w3.org/RDF/ (2004)

9. EMT Madrid Home Page. http://www.emtmadrid.es/ (2012)
10. Service Oriented Architecture (SOA) - W3C Web Open Standards. http://www.w3.org/2008/11/dd-soa.html (2008)
11. Public Regional Consortium of Transports of Madrid Home Page. http://www.ctm-madrid.es/ (2013)
12. Chair of EcoTransport, Technology and Mobility of the Rey Juan Carlos University. http://www.catedraetm.es/
13. Spanish National Society of Blind People. http://www.once.es/ (2013)
14. eXtensible Markup Language (XML) - W3C. http://www.w3.org/XML/ (2013)