

Departamento de Ciências e Tecnologias da Informação

Planeador de Viagens Multimodal

António Gabriel Gomes Saramago Ramos

Dissertação submetida como requisito parcial para obtenção do grau de

Mestre em Engenharia Informática

Orientador(a):

Doutor João Carlos Amaro Ferreira

ISCTE-IUL

Coorientador(a):

Doutor António Ramos Andrade

IST-UTL

Outubro, 2017

Agradecimentos

Aos meus orientadores, Professor Doutor João Ferreira e Professor Doutor António Andrade, agradeço a disponibilidade, orientação e apoio na elaboração da presente dissertação.

À minha esposa e restante família, pela motivação, compreensão e apoio ao longo do meu período académico, no qual, sem a sua ajuda não seria possível conquistar os meus objetivos.

Aos meus colegas, que têm vindo a caminhar comigo nesta vida académica e em especial ao Marco Silva e João Baldeante, agradeço a ajuda, camaradagem e todos os momentos em que o vosso apoio foi fundamental.

À minha entidade empregadora, por ter permitido que conseguisse conciliar o período académico ao longo do tempo.

A todos os meus amigos, agradeço a amizade e principalmente a compreensão pela minha ausência.

Resumo

Cada vez mais nas grandes cidades existem múltiplos meios de transporte sendo que hoje em dia muitos passageiros necessitam de mudar de transporte para se deslocarem. A circulação é frequentemente afetada por acidentes, manutenções na via ou congestionamentos, pelo que o tempo de viagem dos passageiros é aumentado. Consequentemente, os atrasos afetam os horários dos transportes, pelo que, desvios e alterações do meio de transporte, podem ajudar em muito, na redução do tempo de viagem até ao destino.

Os planeadores de viagens multimodais, estão-se a tornar um elemento importante num contexto de *intelligent transport systems*, onde os utilizadores/passageiros podem consultar, planear e escolher a viagem de acordo com os transportes e horários disponíveis. No entanto, o desenvolvimento de sistemas multimodais, onde os dados advêm de diferentes fontes, gera situações e conflitos que carecem de resolução para o correto funcionamento do sistema.

Em alternativa à tradicional solução de integração de dados de diferentes operadores, é apresentada uma solução que permite retirar informação de transportes usando as API disponibilizadas. Estes resultados são agregados num agregador, o qual introduz conteúdo semântico através do uso de uma ontologia de transportes e as diferentes coordenadas geográficas são normalizadas para evitar o problema das diferentes representações de cada operador.

Esta dissertação está inserida num projeto europeu MASAI, no qual faz parte de um conjunto de soluções que permitem aos utilizadores europeus planear as suas viagens. Foram desenvolvidos módulos/serviços que podem ser integrados no projeto MASAI, com informações em tempo-real vindas dos diferentes operadores. Ao longo do trabalho são levantados vários problemas, tendo estes sido investigados e desenvolvidas soluções que permitam reduzir e suprimir os obstáculos na implementação de planeadores de viagens multimodais.

Palavras-Chave: planeador; viagens; multimodal; semântica; API *Mashup*; ontologia; DBSCAN.

Abstract

Increasingly in large cities there are multiple means of transportation and today many passengers need to change transport to move efficiently. Road circulation is often affected by accidents, maintenance or congestion, so travel time of passengers is increased. Consequently, delays affect transport schedules, so transport changes are required and will help reducing the travel time to destination.

Multimodal travel planners are becoming an important element in a context of intelligent transport systems, where users can access, plan and choose the journey according to the transport and schedules available. However, the development of multimodal systems, where data come from different sources, causes conflicts that need resolution for the correct functioning of the system.

As an alternative to the traditional data integration solution of different operators, a solution is presented that allows to retrieve transport information using the APIs available. These results are aggregated in an aggregator, which introduces a semantic content using a transport ontology and the different geographic coordinates are normalized to avoid the problem of the different representations of each operator.

This dissertation is part of a European MASAI project, which is part of a set of solutions that will allow European users to plan their trips. Services/modules have been developed that can be integrated into the MASAI project, with real-time information from different operators. Many problems have been raised throughout this dissertation, where these have been investigated and solutions were developed to reduce and eliminate barriers to the implementation of multimodal travel planners.

Keywords: planner; journey; multi-modal; semantic; API *Mashup*; ontology; DBSCAN.

Índice

Agradecimentos.....	i
Resumo.....	ii
Abstract.....	iii
Índice.....	iv
Índice de Tabelas	vi
Índice de Figuras.....	vii
Lista de Abreviaturas e Siglas	viii
1. Introdução	1
1.1 Enquadramento do tema.....	1
1.2 Motivação e relevância do tema.....	2
1.3 Projeto MASAI	3
1.3.1 Enquadramento do âmbito	4
1.4 Questões e objetivos de investigação	5
1.5 Abordagem metodológica	6
1.6 Estrutura e organização da dissertação.....	8
2. Revisão da Literatura.....	9
2.1 Multimodal Journey Planners	9
2.2 Integração de Dados	10
2.3 Web Semântica.....	14
2.3.1 Ontologia	15
2.3.2 Resource Description Framework	15
2.3.1 SPARQL (SPARQL Protocol and RDF Query Language)	16
2.4 Vehicle Routing Problem	16
2.4.1 Incerteza dos nós da rede	19
2.4.1.1 Clustering k-means.....	20
2.4.1.2 Clustering DBSCAN	21

2.5	API Mashups	22
2.6	Conclusão	23
3.	Multimodal Journey Planner	25
3.1	Visão.....	25
3.2	Arquitetura	26
3.3	Agregador.....	28
3.3.1	Ontologia	30
3.3.2	Implementação	33
3.4	Incerteza dos nós da rede	38
3.5	Conclusão	40
4.	Avaliação do trabalho efetuado.....	41
4.1	Agregador.....	41
4.2	Incerteza dos nós da rede	44
4.3	Conclusão	46
5.	Conclusão	48
5.1	Trabalho Futuro.....	49
	Bibliografia	51
	Anexos e Apêndices.....	56
	Anexo A – Tabela de conceitos da ontologia do Journey Planner.....	57
	Apêndice A – (XML) Exemplo dados do operador X.....	59
	Apêndice B – Exemplo do RDF gerado para o operador X.....	61
	Apêndice C – Exemplo de query para retornar as viagens do ponto Y a partir de Z horas	64
	Apêndice D – (JSON) Exemplo dados do operador Y.....	65
	Apêndice E – Exemplo do RDF gerado para o operador Y	67
	Apêndice F – Exemplo do RDF gerado para o operador X e Y	69

Índice de Tabelas

<i>Tabela 1 – Parte do glossário dos conceitos do PTO no original em inglês (Fonte: Houda, Khemaja, Oliveira, & Abed, 2010).....</i>	<i>32</i>
<i>Tabela 2 – Distribuição de nós ao longo de um determinado espaço</i>	<i>40</i>
<i>Tabela 3 – Parte da informação estruturada do operador X.....</i>	<i>41</i>
<i>Tabela 4 – Parte da informação estruturada do operador Y</i>	<i>42</i>
<i>Tabela 5 – Estrutura única que aglomera os dados do operador X e Y</i>	<i>43</i>
<i>Tabela 6 – Resultados obtidos da query ao agregador.....</i>	<i>43</i>
<i>Tabela 7 – Resultado do DBSCAN na segmentação dos nós por clusters.....</i>	<i>44</i>

Índice de Figuras

<i>Figura 1 – Vista geral do projeto MASAI (Fonte: ASBL, n.d.)</i>	4
<i>Figura 2 – Duas das principais categorias de standards existentes</i>	12
<i>Figura 3 – Relações dos nós entre a web “standard” e a web semântica (Fonte: Wenzel, 2016)</i>	14
<i>Figura 4 – Convergência do clustering k-means em duas iterações sequenciais</i>	20
<i>Figura 5 – Ilustração do modelo de clustering do DBSCAN (Fonte: Schubert, Sander, Ester, Kriegel, & Xu, 2017)</i>	21
<i>Figura 6 – Proposta da arquitetura do sistema</i>	27
<i>Figura 7 – Arquitetura do agregador</i>	29
<i>Figura 8 – Parte da ontologia relacionada com os transportes multimodais (Fonte: Houda et al., 2010)</i>	31
<i>Figura 9 – Parte da ontologia relacionada com o journey planner (Fonte: Houda et al., 2010)</i>	32
<i>Figura 10 – Definição das classes e relações do STOP_POINT</i>	34
<i>Figura 11 - Definição das classes e relações da TRANSPORT_LINE</i>	34
<i>Figura 12 – Excerto de informação recebida pelo operador X (formato XML)</i>	35
<i>Figura 13 – Excerto da informação recebida pelo operador Y (formato JSON)</i>	36
<i>Figura 14 – Excerto da informação guardado dos operadores X e Y no formato RDF</i>	37
<i>Figura 15 – SPARQL query para obtenção de informação dos operadores</i>	38
<i>Figura 16 – Demonstração das posições fornecidas pelos operadores</i>	39
<i>Figura 17 – Representação gráfica dos clusters criados pelo DBSCAN</i>	45
<i>Figura 18 – Representação dos centros dos clusters (a vermelho)</i>	45
<i>Figura 19 – Vista geral do trabalho desenvolvido no contexto do MASAI</i>	46

Lista de Abreviaturas e Siglas

MASAI – *Mobility based on Aggregation of Services and Applications Integration*

EU – *European Union*

ABT - *Account Based Ticketing*

ITS – *Intelligent Transport Systems*

DeMaDIE – *Decision Making in Dynamic Information Environments*

CEN – *European Committee for Standardization*

VRP – *Vehicle Routing Problem*

DBSCAN – *Density-based spatial clustering of applications with noise*

RCSPP – *Resource Constraint Shortest Path Problem*

XML – *Extensible Markup Language*

API – *Application Programming Interface*

JSON – *JavaScript Object Notation*

SOAP – *Simple Object Access Protocol*

REST – *Representational State Transfer*

HTTP – *Hypertext Transfer Protocol*

IFOPT – *Identification of Fixed Objects in Public Transportation*

NeTEx – *Network Timetable Exchange*

SIRI – *Standard Interface for Real-time Information*

UML – *Unified Modeling Language*

WSDL – *Web Service Definition Language*

XSD – *XML Schema Definition*

OTA – *Open Travel Alliance*

PTO – *Public Transportation Ontology*

URI – *Uniform Resource Identifier*

OWL – *Web Ontology Language*

RDF – *Resource Description Framework*

RDFa – *Resource Description Framework in Attributes*

SPARQL – *SPARQL Protocol and RDF Query Language*

W3C – *World Wide Web Consortium*

1. Introdução

1.1 Enquadramento do tema

Nos últimos anos tem havido um elevado crescimento das redes de transporte como forma de ajudar o desenvolvimento das cidades e da sua área periférica permitindo assim responder à exigência dos utilizadores/utentes. Nas grandes cidades existem diferentes modos de transporte, pelo que para o uso eficiente e sustentável na deslocação dos passageiros, pode ser necessário a alteração de transporte ao longo do percurso do passageiro.

A utilização de transportes públicos em grandes cidades é largamente utilizada, sendo que um dos fatores que contribui para o aumento da qualidade da viagem do utente é a informação facultada pelo transporte público (por exemplo, o tempo que este demora a chegar) e a possibilidade de saber como posso ir para o local X, tendo em conta a sua posição atual e a diversidade de transportes.

O facto de muitos passageiros necessitarem de vários transportes ou mudarem de operador para trajetos pontuais, aumenta a complexidade e duração no planeamento da viagem. Um planeador de viagens multimodal irá permitir responder a um utilizador através de um ou vários serviços de transporte, qual o melhor caminho a efetuar de um determinado local para outro a uma determinada hora. Este permitirá a junção de diversos níveis de operadores (municipais, nacionais e até internacionais).

Pretende-se investigar soluções para os problemas científicos que se levantam na implementação de soluções que permitam determinar a melhor forma de um utilizador se deslocar numa cidade tendo em conta a sua posição atual e os transportes públicos disponíveis.

1.2 Motivação e relevância do tema

Os planeadores de viagem com informação em tempo real são aplicações importantes para, num contexto de cidades inteligentes, as pessoas usarem cada vez mais os transportes públicos em detrimento do transporte individual e assim contribuir para a mitigação de emissões de CO₂. Este facto, associado ao crescente uso dos dispositivos móveis no dia-a-dia, constitui uma oportunidade para o desenvolvimento de uma solução que permita a melhor escolha de transportes tendo em conta a diversidade de operadores existentes e as preferências dos utilizadores.

Ao longo dos anos, alguns dos problemas têm vindo a ser investigados (Rocha, 2012) (Haghani & Jung, 2005) (Borole, Rout, Goel, Vedagiri, & Mathew, 2013) (Gkiotsalitis & Stathopoulos, 2015), no entanto, estes são mais centrados para a resolução dos problemas de *routing*, sendo que a sua aplicação prática associada a dispositivos móveis e a normalização dos dados entre todos os operadores não é abrangida.

Este tem sido um tópico interessante de investigação onde vários problemas científicos podem ser estudados e aprofundados:

1. Integração da informação dos diversos operadores, tendo em conta a diversidade dos dados;
2. Adaptação de algoritmos de *routing*, resolução de conflitos geográficos dos nós e a apresentação contextualizada dos outputs;
3. Personalização da informação tendo em conta a posição geográfica;
4. Problemas de usabilidade dado as dimensões reduzidas de apresentação de informação;
5. Problemas da interface quando existe necessidade de mudar de operador de transporte público.

O objetivo da dissertação foi o estudo de soluções para o problema da normalização dos dados, a sua interoperabilidade entre os diferentes operadores e a resolução de conflitos das posições geográficas provenientes dos vários operadores.

1.3 Projeto MASAI

A indústria de viagens e turismo está em mudança derivado à evolução dos requisitos dos consumidores e às novas possibilidades que a tecnologia nos trouxe. Desta forma, estão a surgir novas oportunidades de negócios em que o custo é menor comparativamente aos anos anteriores.

Implementado sobre o programa EU HORIZON 2020, o *Mobility based on Aggregation of Services and Applications Integration* (MASAI)¹ é um projeto de pesquisa e desenvolvimento para criar um ecossistema aberto, que se insere no ambiente digital na área de transportes. A visão do MASAI é satisfazer as necessidades gerais de serviço de transportes exigidas pelos “cidadãos móveis” a “agregação de funcionalidades de viagem no seu bolso”(ASBL, 2016).

O MASAI pretende alcançar o objetivo, através da constituição de uma comunidade de partes interessadas que contribuam continuamente para o desenvolvimento dos elementos centrais de um ecossistema digital. Desta forma e como retratado na Figura 1, é permitido através da arquitetura proposta, constantes melhorias e a sustentação de futuros desenvolvimentos técnicos.

Assim, o projeto MASAI visa definir, implementar e testar uma arquitetura que possui as seguintes características: 1) Desmaterialização de bilheteiras garantindo segurança na utilização de software de bilheteira; 2) Integração com pagamentos bancários e comissões; 3) Suporte de *Logical Account Based Ticketing* (ABT); 4) Baseado num conceito modular e com interfaces bem definidas a publicação através de *API's open-source* para facilitar a integração com dispositivos móveis; 5) Fornecer mecanismos de estatísticas e *report* contínuo (*check-in*, *check-out*, validações, etc...); e 6) Compatibilidade com formatos e estruturas já existentes (ASBL, 2016).

¹ <http://masai.solutions>

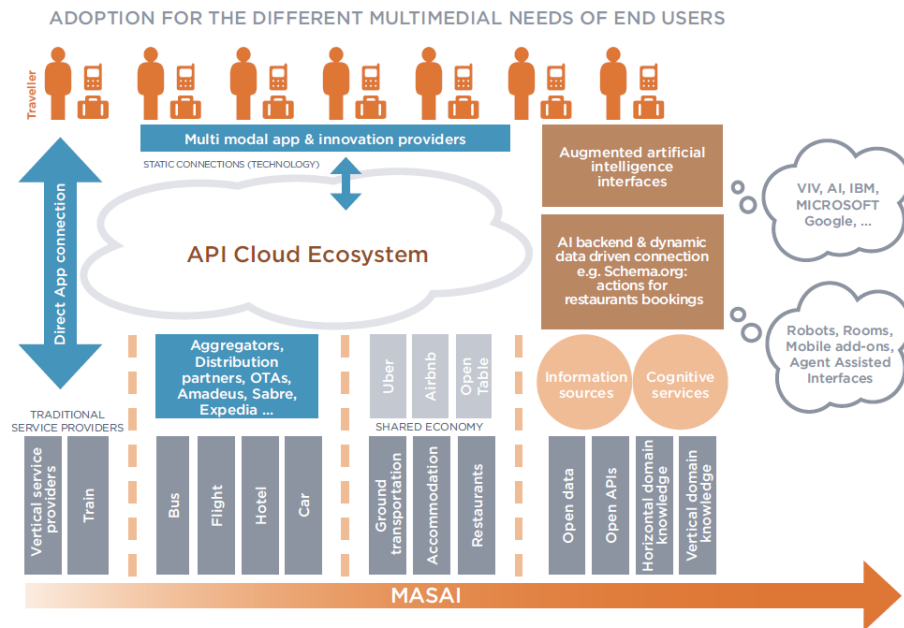


Figura 1 – Vista geral do projeto MASAI (Fonte: ASBL, n.d.)

1.3.1 Enquadramento do âmbito

No âmbito deste trabalho será formada uma proposta e desenvolvida uma prova de conceito para o quarto objetivo do projeto MASAI. Este objetivo foi inicialmente definido em parceria com uma empresa externa, em que esta efetuava a ponte com os operadores e fornecia através de *API's* a informação dos operadores de transporte.

No âmbito da dissertação foram estudados e avaliados métodos, formatos e tecnologias de integração de dados com o propósito de garantir a normalização da informação dos operadores de transporte. Foi também sugerida uma arquitetura que permita concretizar as operações necessárias para o quarto objetivo do projeto MASAI, tendo em conta as suas características e particularidades. O estudo do VRP (*Vehicle Routing Problem*) e de algoritmos de *routing*, é efetuado no sentido de lidar com o problema da normalização da informação dos diferentes operadores. Este necessita de ser resolvido para os algoritmos de *routing* poderem ser aplicados sobre os dados normalizados e, posteriormente no contexto do projeto MASAI, o algoritmo devolver o melhor percurso do ponto A ao ponto B.

A prova de conceito foi segmentada em diferentes componentes que conjugados permitem responder ao quarto ponto proposto pelo MASAI. Os componentes a serem

abordados são: 1) arquitetura de um sistema central; 2) Integração dos diferentes operadores; 3) Resolução dos conflitos e normalização dos dados geográficos.

1.4 Questões e objetivos de investigação

A integração de dados em sistemas de transporte tem sido uma preocupação do CEN (*European Committee for Standardization*) pelo que a investigação dos modelos, protocolos, tecnologias existentes e como estes se enquadram, é necessária para avaliar a sua aplicação no contexto do planeador multimodal.

A centralização e a agregação dos dados tornam-se essenciais para o sucesso deste tipo de aplicações. O presente trabalho insere-se no contexto do projeto MASAI, contribuindo com o agregador de dados, a normalização dos dados e das coordenadas geográficas.

A aplicação de algoritmos de *routing* em sistemas multimodais, levanta problemas quando a informação devolvida pelos vários operadores não é normalizada. Caso este problema não seja minimizado, gera incerteza no cálculo da melhor rota, inviabilizando a qualidade e precisão do algoritmo de *routing*. Para esse efeito, foi desenvolvida uma aproximação de *clustering* geográfico de forma a reduzir as coordenadas espaciais das paragens a uma só.

Num contexto semântico, torna-se necessário a utilização de uma ontologia para unir e normalizar os objetos retornados pelos diferentes operadores. A ontologia deve ser abrangente o suficiente, pois ao ser empregue aos diferentes meios de transporte esta deverá manter o significado dos termos e vocabulários.

A integração dos sistemas, a agregação dos operadores e a resolução dos problemas inerentes são os objetos de estudo neste trabalho. São identificados protocolos e ferramentas que permitam a normalização dos dados e a sua interoperabilidade; é exposto o problema na consistência dos dados, sendo sugeridos métodos de correção da informação para evitar conflitos no cálculo do caminho mais curto (*Shortest Path*) consoante as restrições existentes; identifica-se e é utilizada uma ontologia que permita a integração semântica dos dados; é proposto uma arquitetura central capaz de integrar os dados multioperador.

Assim o presente trabalho inclui os seguintes objectos alvos de investigação: 1) integração de dados, tendo em conta a heterogeneidade das bases de dados dos diferentes operadores de transportes públicos; 2) normalização e resolução de conflitos de nós com *clustering*, permitindo a integração da informação eficazmente no algoritmo de *routing*; 3) implementação de um conjunto de componentes que integram a arquitetura de um sistema central, no qual será aplicado no projeto MASAI².

1.5 Abordagem metodológica

Como método de investigação, foi utilizado para este trabalho o DSR (*Design Science Research*) pois permite com base no paradigma da resolução de problemas estender os limites humanos e das organizações criando novos e inovadores artefactos.

O objetivo de definir um processo de DSR foi fornecer um modelo para a elaboração do resultado da pesquisa. Os autores (Peffer, Tuunanen, Rothenberger, & Chatterjee, 2007) sugerem a utilização de um modelo com seis atividades sequenciais:

- Identificação do problema e motivação
- Definição dos objetivos da solução
- Desenho e desenvolvimento
- Demonstração
- Avaliação
- Comunicação

No entanto e dependendo da forma como a investigação é criada, esta pode ser iniciada a partir de um diferente ponto. No contexto deste trabalho foram seguidos e detalhados os processos nos passos que os autores referem.

• Identificação do problema e motivação

A conceção de um planeador de viagens multimodal levanta problemas na integração dos dados de diversos operadores, pois cada um, utiliza um modelo de dados diferente e a sua interoperabilidade fica condicionada, não existindo uniformidade nos dados

² A integração dos componentes no sistema central do MASAI não está inserida no âmbito do projeto.

(horários, rotas, paragens, etc...); resolução de conflitos e redução da incerteza dos nós; um sistema central irá permitir efetuar operações tendo em conta os diferentes tipos de dados.

- **Objetivos da solução**

O estudo e as propostas de soluções para os problemas, foram o objetivo primordial deste trabalho. Em primeiro lugar, é necessário pesquisar e identificar ferramentas que permitam a normalização dos dados e a sua interoperabilidade. Em segundo, é necessário identificar os problemas de *routing* no cálculo do caminho mais curto (*Shortest Path*) consoante as restrições existentes. Em terceiro, uma arquitetura central capaz de integrar os dados multioperador e consequentemente com a capacidade de achar o caminho ideal para percorrer de um ponto X a Y.

- **Conceção e desenvolvimento**

Nesta atividade foi desenvolvida uma prova de conceito que demonstra a aplicação das sugestões propostas aos problemas acima descritos. Estes foram segmentados em componentes e o processo que levou à concretização do artefacto encontra-se descrito ao longo do documento.

- **Demonstração**

Nesta fase foram utilizados os componentes implementados, com o objetivo de analisar se estes permitem resolver os problemas propostos e se podem ser integrados numa arquitetura multimodal.

- **Avaliação**

Na fase de avaliação da prova de conceito, foi possível verificar se os resultados permitem resolver os problemas de agregação, integração e normalização dos sistemas e dos dados.

- **Comunicação**

Parte do contributo deste trabalho foi difundido nas redes académicas, através de um artigo desenvolvido em conjunto com outros autores (Ferreira, Andrade, Ramos, Martins, & Almeida, 2017), onde foi integrado no livro (*Communications in Computer and*

Information Science) publicado pela Springer e apresentado no DeMaDIE (*Workshop on Decision Making in Dynamic Information Environments*).

1.6 Estrutura e organização da dissertação

A presente dissertação está organizada em cinco capítulos que pretendem refletir as diferentes fases até à sua conclusão.

O primeiro capítulo introduz o tema, motivação e relevância da investigação, enquadra o tema no projeto europeu MASAI, define os objetivos, questões de investigação e refere uma breve descrição da estrutura do trabalho.

O segundo capítulo, designado por Revisão da Literatura, reflete o enquadramento teórico do tema, sendo abordados os tópicos investigados e os problemas existentes da sua aplicação no contexto do projeto.

O terceiro capítulo é dedicado ao *Multimodal Journey Planner*, sendo especificado a visão, a arquitetura proposta e a implementação dos componentes/serviços.

No quarto capítulo é apresentado a avaliação dos resultados obtidos pelos vários componentes que compõem a prova de conceito.

No quinto e último capítulo apresentam-se as conclusões da dissertação, bem como as limitações existentes e trabalho futuro.

2. Revisão da Literatura

Os temas de investigação estão relacionados com três pontos principais de investigação: 1) a integração de dados de diferentes operadores, com o problema inerente da existência de diversos formatos; 2) Normalização e resolução dos conflitos da informação para aplicação no algoritmo de *routing*; 3) arquitetura e implementação do agregador da informação dos operadores.

Na conceção da prova de conceito, a integração e normalização dos dados multioperador é crucial para o correto cálculo da melhor rota pelo sistema. Sem isso, o algoritmo de *routing* perderia performance podendo até mesmo ser impossível calcular a rota.

Além destes três grandes problemas que constituem parte crucial do trabalho de investigação, outros foram detetados e serão discutidos nos capítulos seguintes.

2.1 *Multimodal Journey Planners*

Com os mais recentes avanços na tecnologia, agregar e integrar a informação dos diferentes tipos de transporte está a tornar-se um fator importante para o desenvolvimento de Sistemas de Transporte Inteligentes (ITS).

Através da informação estática obtida dos diversos operadores é permitido ao utilizador obter informação do percurso, os horários e paragens de cada percurso. Assim e em casos onde seja necessário circular em diferentes transportes, a eficiência na obtenção da melhor e mais rápida rota, pode-se tornar num processo exaustivo e penoso.

No contexto empresarial, a informação que é facultada pelo operador adota uma estrutura de dados proprietária que acaba por ser favorável ao contexto da empresa. Assim a sua conjugação com outros operadores/fornecedores pode ficar inviabilizada no futuro.

Os planeadores multimodais são sistemas de agregação que permitem responder a questões de como se deslocar de um ponto X a um ponto Y, a partir de uma hora específica, utilizando um ou múltiplos operadores. Estas ferramentas possibilitam o

aumento da eficácia, pois combinam a informação dos múltiplos operadores, em apenas um sistema, com o objetivo de obter toda a informação dos múltiplos operadores para o contexto que o passageiro necessita.

No capítulo 3.1 descreve-se um conjunto de problemas na implementação destes sistemas, sendo que no capítulo seguinte é descrito o problema da diversificação da informação disponível e a integração dos operadores para uma correta compreensão pelo cliente (Ferreira, Andrade, Ramos, Martins, & Almeida, 2017).

Além deste problema, o cálculo da melhor rota continua a ser um importante problema num contexto multimodal. Foram desenvolvidos estudos sobre o *multimodal shortest viable path problem* (Ziliaskopoulos & Wardell, 2000) (Lozano & Storchi, 2001) (D. López & Lozano, 2014) onde são apresentadas técnicas de pré-computação dos caminhos possíveis, algoritmos de *routing*, utilização de restrições e sequências dos nós. No capítulo 2.4.1 é aprofundado esse tema, sendo apresentado um problema que surge na integração de múltiplos operadores onde subsistem conflitos de dados.

2.2 Integração de Dados

O armazenamento dos dados adquiridos através do conhecimento do ser humano tem vindo a evoluir ao longo do tempo. Desde o uso de pergaminhos, bibliotecas, bases de dados, repositórios digitais, os dados são continuamente guardados para posterior utilização e enriquecimento de conhecimento.

Nos nossos dias e no contexto das novas tecnologias a busca de informação específica utiliza mecanismos que favorecem e que facilitam o acesso a dados de diversas entidades externas. Podemos tomar como exemplo as aplicações móveis que transmitem a meteorologia. Existem inúmeras entidades meteorológicas que facultam dados para extração e consumo público, mas a forma de recolher os dados de cada uma destas pode forçar a utilização de um método e/ou uma estrutura de dados diferente.

Hoje em dia com o propósito de recolher e integrar dados de sistemas de informação externos, podem ser utilizados *web services* que retornam dados com base em *web standards* (Ex: XML, SOAP, REST) e que operam principalmente sobre protocolos de

HTTP. Apesar da utilização de *web services* permitir um desenvolvimento mais rápido e económico, ao escalarmos a informação recebida para dados em tempo real e o uso de multioperadores, a agregação dos dados torna-se muito complexa sendo que o uso da mesma estrutura para os diversos operadores não irá garantir o correto armazenamento dos mesmos.

A União Europeia e o Comité Europeu para a Standardização têm vindo a efetuar uma caminhada para a definição de regras, guias e padrões para a distribuição de informações multimodais (Standardization, 2010).

Estes já definiram um conjunto de *standards*, nomeadamente: *TransModel* que fornece uma *framework* para definição de modelo conceptual de dados (Standardization, 2015); o IFOPT que é um *standard* que define um modelo de dados para a identificação de elementos físicos do transporte público (paragens, estações, percursos, etc...) (CEN/TC 278, 2008); o NeTEx é um complemento ao *TransModel* que define um *standard* técnico para o intercâmbio de horários de transportes públicos e a sua informação relacionada (Arneodo, 2015); o SIRI (*Standard Interface for Real-time Information*) é um *standard* para o intercâmbio de dados em tempo real (Bruns, 2010); o DATEX II foi desenvolvido com o objetivo de trocar informações de trânsito entre os diversos sistemas de informação (Standardization, 2013).

Na Figura 2 é apresentada uma ilustração da segmentação das duas principais categorias dos *standards* já existentes. Estes são divididos em dois principais grupos com funções distintas: *Model Data* – Permite a indexação, organização e relacionamento das várias informações; *Data Exchange Interfaces* – Define os protocolos e formatos usados na transmissão de dados.



Figura 2 – Duas das principais categorias de standards existentes

O *TransModel* fornece uma *framework* que possibilita a definição de modelos de dados para toda a rede pública de transporte (Standardization, 2015). A adoção de *standards* melhora a interoperabilidade entre os vários operadores de transporte, pois estes acabam por utilizar as mesmas definições e estruturas garantindo assim a semelhança dos diversos sistemas na mesma solução.

O IFOPT (*Identification of fixed objects in public transportation*) é uma especificação técnica do CEN que define um modelo de dados para a identificação de objetivos fixos na rede pública de transportes, por exemplo, estações, paragens de autocarros, aeroportos, portos náuticos. É possível enquadrar os objetos nas seguintes categorias (CEN/TC 278, 2008):

- Objetos fixos:
 - Paragens de autocarro;
 - Pontos de acesso;
 - Pontos de interesse.
- Objetos móveis:
 - Veículos.
- Eventos:
 - Acidentes;
 - Congestionamento.

O NeTeX define um *standard* para a troca de informação de passageiros. A informação é codificada em documentos XML que deve corresponder a um *XML Schema*. Estes podem ser transmitidos via HTTP ou FTP. Este foi concebido para trabalhar com os *standards* do modelo de dados do *TransModel* sendo que ainda permite a interação com o IFOPT e o SIRI, garantindo assim a compatibilidade e a interoperabilidade entre os diversos *standards* (Arneodo, 2015).

O SIRI (*Service Interface for Real-Time Information*) é o *standard* criado pelo CEN para a troca de informação em tempo real entre os vários sistemas de informação. A informação é transacionada em formato XML e usa o *TransModel* como base dos modelos de dados. Este foi desenvolvido com o apoio de várias entidades espalhadas pela Europa e como tal, incorpora propriedades e *standards* definidos conjuntamente pelos vários países e organizações.

O DATEX II é um *standard* criado e mantido pelo CEN, tendo sido desenvolvido com o objetivo de trocar informação de trânsito entre os centros de gestão de tráfego, centros de informação de tráfego e fornecedores de serviço (Standardization, 2013). A especificação do DATEX II permite uma normalização entre toda a informação gerada e processada, sendo que a nível europeu esta garante uma maior uniformidade entre todos os parceiros, operadores e sistemas de controlo. O DATEX II especifica a sua estrutura de modelo de dados em UML (*Unified Modeling Language*) sendo que este atua através de XML, podendo ser depois transmitida através do protocolo HTTP e/ou *webservices* (WSDL ou SOAP).

Contudo e apesar dos esforços da União Europeia, os operadores e prestadores de serviço ainda não estruturaram os seus dados nos formatos pré-estabelecidos. Alfonso et al., (2016) referem que as utilizações das normas/*standards* do CEN não são claras e que em alguns casos, estes não respondem às restrições e requisitos das cidades.

Os formatos que os operadores de transporte suportam, têm vindo a ser desenvolvidos individualmente, pelo que os sistemas integrados de transporte, para garantir um correto funcionamento acabam por ser obrigados a gerir diferentes *datasets*. Porém, no momento, ainda não existe com sucesso uma especificação, que permita efetuar referências cruzadas dos diferentes dados (Alfonso et al., 2016).

Assim, no contexto desta dissertação foi adotado para o sistema central (agregador), um contexto semântico que permite corresponder os mesmos termos/vocabulário às diferentes fontes e estrutura de dados. Esta opção garante que os dados de diferentes origens respeitem uma ontologia, a extração semântica da informação e a integração com sistemas de *Open Data* ou *Linked Data*.

2.3 Web Semântica

O conteúdo é facilmente legível pelos humanos, mas ao contrário, os sistemas computadorizados sem contexto semântico não têm essa capacidade. O conceito da web semântica foi introduzida em 2001 por Tim Berners-Lee e tem vindo a ser utilizada para aperfeiçoar os motores de busca e criar uma maior disseminação da informação (Dotsika, 2010).

Na Figura 3 são demonstradas as diferenças dos tipos de informação e das relações entre a web “standard” e a web semântica. Enquanto que na web “standard” os documentos podem estar ligados entre si, no contexto da web semântica, um conjunto de nós estão relacionados entre si, onde contêm propriedades e as suas relações contêm um significado que define a sua relação (Wenzel, 2016).

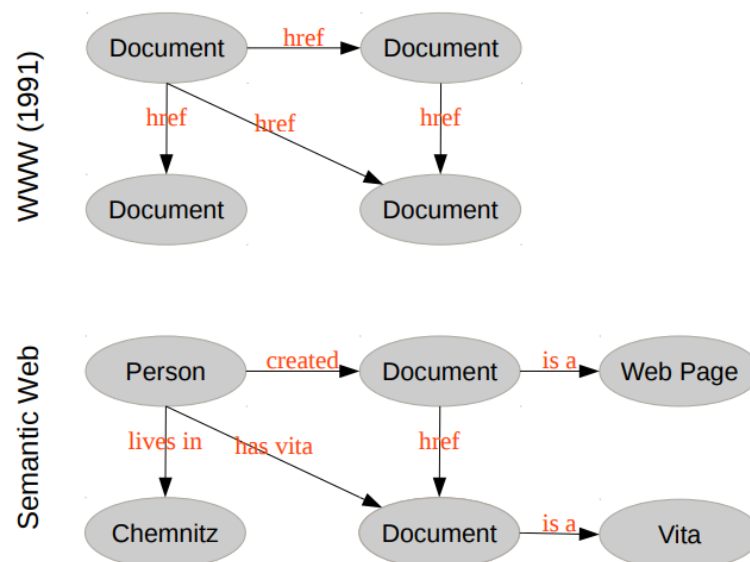


Figura 3 – Relações dos nós entre a web “standard” e a web semântica (Fonte: Wenzel, 2016)

2.3.1 Ontologia

No contexto de web semântica, o modelo do domínio define e descreve os conceitos, relações, entidades e as suas propriedades dentro do domínio (Houda, Khemaja, Oliveira, & Abed, 2010) (Taheriyani, Knoblock, Szekely, & Ambite, 2014). O primeiro passo na criação de uma ontologia é a especificação das classes, propriedades e a classificação de cada informação. O passo seguinte é a especificação das relações entre os atributos dos dados de origem aos termos da ontologia, no entanto, para a sua aplicação em processos automáticos é necessário estender e criar processos que o permitiam. Estudos recentes (Lee & Kim, 2012b) (Lee, 2014) (Taheriyani et al., 2014) têm vindo a aprofundar a utilização de mecanismos automáticos para a descoberta e mapeamento automático de *API's* utilizando as suas descrições semânticas.

Criado pelo *OWL Working Group*³, a *W3C Web Ontology Language* (OWL) é uma linguagem semântica para representar objetos ricos e complexos, grupos de objetos e as relações entre os mesmos (OWL Working Group, 2012). O OWL é utilizado para processar a informação permitindo adicionar vocabulário de forma a facilitar a interpretação do conteúdo XML ou RDF.

O *Linked Data* é um conjunto de boas práticas para conectar e publicar informação estruturada na web. A sua utilização tem vindo a criar um espaço global de dados que contem milhões de afirmações. Em suma, este permite a criação de ligações semânticas entre as várias informações de diferentes fontes de dados. O *Linked Data* utiliza o RDF para efetuar afirmações que ligam a termos/significados (Bizer, Heath, & Berners-Lee, 2009), principalmente sobre duas tecnologias (URIs e HTTP), fornecendo significado a uma entidade a partir de um formato mais genérico.

2.3.2 Resource Description Framework

O RDF (*Resource Description Framework*) é um método *standard* de anotação semântica, que adiciona funcionalidades que permitem a modelação da informação. Esta

³ <https://www.w3.org/2001/sw/#owl>

é efetuada com expressões dos termos sujeito-predicado-objeto, fornecendo significado a termos ao representar e identificar conjuntos de informação em *URI's*.

Os links RDF formam *RDF Triples*, onde o sujeito do *triple* é uma referência URI ao *namespace* de um *dataset*, enquanto o objeto do *triple* é uma referência URI ao outro (Bizer, Heath, & Berners-Lee, 2009). Embora cada pedaço de informação seja particularmente pequeno, a sua integração dentro de uma rede de Web Data pode ser vista, como um exemplo de *Big Semantic Data*. (Cáceres, Sierra-Alonso, Cuesta, Vela, & Cavero, 2015)

O RDF tem vindo a ser adotado em diversos campos de investigação levando à sua expansão tecnológica. O projeto *Linked Open Data* tem um papel crucial na sua evolução ao encorajar a publicação de dados semânticos, ao fornecer identidade aos recursos usando *URIs* (Cáceres et al., 2015).

2.3.1 SPARQL (*SPARQL Protocol and RDF Query Language*)

O *SPARQL* é a linguagem recomendada pela W3C para efetuar *queries* ao RDF. Este pode ser usado em múltiplas fontes de dados, sejam estes guardados nativamente em RDF ou visualizados por intermédio de outro *software*. O *SPARQL* permite efetuar *queries* a padrões de grafos (sejam estes obrigatórios ou opcionais) ao longo das suas junções e disjunções. O resultado das *queries* podem ser conjuntos de resultados ou então a geração de outros grafos RDF (W3.org, 2007).

2.4 Vehicle Routing Problem

Hoje em dia a exigência dos utilizadores e a sua satisfação é bastante importante para o sucesso de qualquer avanço tecnológico. A geolocalização, o sistema de posicionamento global (GPS) e os algoritmos de *routing* têm vindo a evoluir ao longo dos anos e tornam-se numa característica crucial para sistemas de navegação.

A participação de entidades externas torna-se obrigatório, na medida que é necessária a sua contribuição com dados e informações, nomeadamente: horários, rotas, restrições e informação de acidentes. Guardar e manter estes registos invoca a necessidade de um

servidor central que permita a interação entre os dispositivos móveis e um repositório de informação.

Gkiotsalitis & Stathopoulos, 2015, referem que a utilização de dispositivos móveis facilita a interação entre a central e os dispositivos, podendo os mesmos ainda contribuir para o fornecimento de informação em tempo real. Estes mencionam que a grande desvantagem da utilização dos dispositivos móveis para planeador de viagens é a sua capacidade limitada de alojamento e o seu fraco poder de processamento. Por isso, a utilização de uma arquitetura centralizada vem mitigar estas fraquezas sendo que ainda possibilita a disseminação da informação em tempo real para os utilizadores.

A integração e o conjunto da informação dos operadores são cruciais para o cálculo da melhor rota, pelo que o estudo do *Vehicle Routing Problem*, permite compreender a importância deste tema no contexto do projeto MASAI. Além disso, a incerteza dos pontos geográficos (nós), necessita de ser previamente resolvida antes dos dados serem armazenados para depois aplicar o algoritmo de *routing*.

O problema de *routing* de veículos é um problema fundamental na determinação da melhor rota tendo em conta algumas limitações. Quando o problema começou a surgir numa fase inicial eram implementados algoritmos tradicionais (Ex: *Dijkstra* e *A**) para encontrar o *Shortest Path* mas atualmente e no contexto de algoritmos dinâmicos (com dados em tempo real) a sua utilização não é eficiente (Rocha, 2012). Nos problemas com informação em tempo real o objetivo de aumento de performance só se consegue com a conjugação de diferentes cálculos (Ghiani, Guerriero, Laporte, & Musmanno, 2003).

Gkiotsalitis e Stathopoulos referem estudos de autores que fornecem algoritmos para aumentar a eficiência dos algoritmos *Shortest Path* em redes multimodais, sendo que na sua maioria utilizam *labels* associados aos nós, sendo que estes incorporam medidas de correção das suas *labels* à medida que os problemas ocorram (Ex: acidente, congestionamento). Segundo Rocha, uma abordagem baseada no pré-processamento hierárquico apresenta melhores resultados de cálculo (eficiência), quando o número de nós e de arestas que constituem a topologia aumenta (Rocha, 2012).

O *routing hierárquico* é uma técnica que junta três algoritmos onde é efetuado um pré-processamento permitindo assim acelerar o cálculo de rotas (Rocha, 2012). Rocha no âmbito do seu trabalho utiliza o algoritmo *Highway Hierarchies* para efetuar a classificação do percurso (estradas), efetuando uma pesquisa bidirecional ignorando percursos de menor importância; o algoritmo *Highway-node routing* que permite a reação a situações inesperadas; e o algoritmo *Transit-node routing* para permitir encontrar grafos específicos e calcular as distâncias entre os mesmos.

Gkiotsalitis & Stathopoulos (2015), apresentam um algoritmo com uma abordagem heurística e uma taxa de certeza entre os 90% e os 100% e que pode diferir cerca de 3 minutos para viagens curtas, (cerca de 30 minutos) a 9 minutos para viagens superiores a 90 minutos. Estes sugerem uma solução para o grande desafio que é calcular a melhor rota com as limitações existentes. Sugerem uma solução que consiste na definição de um custo (que ilustra o nível de preferência de uma característica da rota do utilizador) em cada arco (conexão entre nós) e executar o algoritmo *Dijkstra* como nos métodos tradicionais. O utilizador deverá definir um limite máximo para cada característica (Ex.: efetuar X mudanças de transporte; Possibilidade de andar Y metros). Esta “solução” é um problema denominado *Resource Constraint Shortest Path Problem* (RCSP) e é *NP-Hard Problem* (Gkiotsalitis & Stathopoulos, 2015) na medida que a complexidade é gravemente afetada pelos modelos de dados dos horários e mudanças de transporte público, e não é possível determinar uma solução ótima em tempo polinomial.

Os autores referem e optam por um conceito que ajuda a solucionar o problema (utilizado por outras entidades, nomeadamente, Google Navigation, Bing Maps). O conceito consiste em procurar todos os caminhos eficientes com a premissa que o objetivo não pode ser otimizado sem piorar o outro. Para ajudar a solucionar os múltiplos objetivos no caminho mais curto (*Shortest Path*), pode ser feita uma procura extensa na rede de nós, sendo facilitada por algoritmos que criam, ajustam e corrigem as *labels* dos nós. Como o utilizador tem selecionado uma preferência predominante, o caminho escolhido é o mais próximo dentro dos limites definidos por este (Gkiotsalitis & Stathopoulos, 2015).

Tanto Rocha como Gkiotsalitis & Stathopoulos (2015), com o objetivo de otimizar o cálculo do caminho mais curto propõem a utilização de algoritmos de pesquisa que permitem refinar e eliminar caminhos ineficientes.

No contexto deste trabalho, foi aprofundada a questão da incerteza dos nós da rede, pois este torna-se um grande problema em sistemas multimodais, quando se integra informação dos diferentes operadores.

2.4.1 Incerteza dos nós da rede

As estações e paragens (nós da rede) são componentes das linhas dos transportes públicos, no entanto, o mesmo nó pode ser associado a diferentes linhas e até a diferentes operadores.

Jain, Jain, & Parida (2014) referem que o maior entrave para os planeadores multimodais é a falta de integração da informação de diferentes fontes. Além disso, os autores referem ainda que as informações disponíveis dos operadores contêm diferenças na nomenclatura das paragens. Ao longo deste trabalho, foi detetado um outro problema, associado ao algoritmo de *routing* e à incerteza na localização das estações/paragens (nós) da rede, nomeadamente nos transbordos entre operadores de redes diferentes.

Como tal, além da nomenclatura, mas maioritariamente a incerteza da localização das paragens (nós), dificulta o cálculo do algoritmo de *routing*. Um procedimento simples para abordar/atacar este problema de incerteza nas localizações de estações e paragens dos diferentes modos de transporte, seria apenas permitir transbordos se as estações/paragens estivessem dentro de uma determinada distância ($\Delta\delta$). Assim, se a distância entre duas estações for menor ou igual a $\Delta\delta$, então o transbordo é permitido, caso contrário não é.

Uma abordagem possível para a solução deste problema passa pela aplicação de algoritmos de *clustering* ao conjunto de nós, para inferir grupos de nós que pertençam à mesma paragem/estação.

Posteriormente, é possível refinar este procedimento, fazendo depender esta quantidade limite $\Delta\delta$ do par de modos de transporte entre os quais ocorre o transbordo. Por exemplo, para transbordos entre metro e aeroporto, o valor $\Delta\delta$ assumiria valores maiores do que transbordos entre autocarros (muitas vezes na mesma paragem).

2.4.1.1 Clustering k-means

O algoritmo de *clustering k-means* é um algoritmo baseado na distância. Este usa a distância como método de avaliação, onde N nós, são divididos em K conjuntos compactos e independentes. A distância assume formas *hyperspherical* ou *hyperelliptical* individuais para agrupar os resultados (Ben Khediri, Weihs, & Limam, 2012).

A Figura 4 exemplifica duas iterações durante a execução do algoritmo *k-means* e a definição do conjunto em tom de azul. Ao longo das iterações, os nós são classificados para o conjunto mais perto que se encontra o nó e o centro do conjunto. Em cada iteração, o centro do conjunto é recalculado e o processo é repetido até haver convergência.

A utilização deste algoritmo para o *clustering* dos nós, iria seleccionar os pontos iniciais do conjunto aleatoriamente sendo que não permite uma correta identificação e separação dos conjuntos.

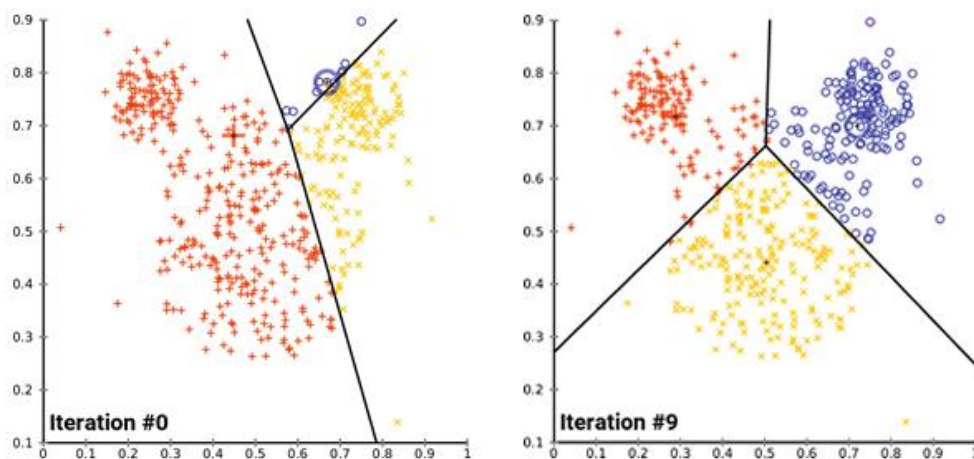


Figura 4 – Convergência do clustering k-means em duas iterações sequenciais⁴

⁴ Adaptado de Wikimedia Commons

2.4.1.2 Clustering DBSCAN

O *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) é um algoritmo de *clustering* baseado na densidade dos pontos (*density-based clustering algorithm*). Este permite através de um conjunto de pontos espalhados num determinado espaço, agrupá-los num único conjunto.

O objetivo do algoritmo é a descoberta destes conjuntos/áreas, que satisfaçam a densidade mínima, e nos quais são separados pelas áreas de baixa densidade. O modelo introduzido pelo DBSCAN, usa uma estimativa mínima do nível da densidade baseada no *threshold* do número dos vizinhos, *minPts*, dentro do raio ϵ (com uma distância arbitrária). Objetos com mais de X (parâmetro *minPts*) vizinhos dentro do raio são considerados como um conjunto (Schubert, Sander, Ester, Kriegel, & Xu, 2017).

Através da Figura 5 é possível verificar uma ilustração dos conceitos do DBSCAN, onde o parâmetro *minPts* foi definido igual a quatro e o raio ϵ está definido pelos círculos. O ponto N, é um “ponto ruído”, o ponto A um ponto chave e os pontos B e C são “pontos fronteira”. As setas definem o alcance da densidade, sendo que os pontos B e C estão ligados pela densidade pois ambos são alcançados pela densidade do ponto A. Como o ponto N não é alcançado pelo ponto A (ponto chave) este é considerado como um “ponto ruído” (Dudik, Kurosu, Coyle, & Sejdić, 2015) (Schubert, Sander, Ester, Kriegel, & Xu, 2017).

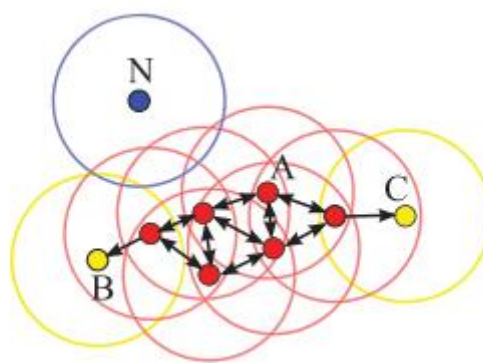


Figura 5 – Ilustração do modelo de clustering do DBSCAN (Fonte: Schubert, Sander, Ester, Kriegel, & Xu, 2017)

O algoritmo percorre circularmente todos os pontos do conjunto e calcula o número de vizinhos que cada ponto possui. Os conjuntos são definidos com o número de outros pontos que estão dentro da distância mínima do ponto original. Qualquer ponto que não

tenha o número mínimo de vizinhos (definido no parâmetro *minPts*) é declarado um ponto de ruído e não é associado a um conjunto (Dudik, Kurosu, Coyle, & Sejdić, 2015).

O DBSCAN tem similaridades computacionais comparativamente ao *k-means*, no entanto, ao contrário do *k-means*, o DBSCAN utiliza a densidade dos pontos pelo espaço para identificar conjuntos invés de identificar a localização dos pontos centrais.

Dudik, Kurosu, Coyle & Sejdić (2015) referem que as vantagens da utilização do DBSCAN são: 1) superior identificação e separação dos conjuntos de diferentes tamanhos e formatos; 2) ordenação dos pontos de ruído para um conjunto separado, caso o ponto seja similar ao restante *dataset*; 3) é determinístico, ao contrário do *k-means* que aleatoriamente seleciona os pontos iniciais do conjunto.

2.5 API Mashups

Como discutido anteriormente, a normalização e a interoperabilidade dos dados é um dos pontos críticos para os planeadores multimodais. Porém um sistema que permita a utilização de várias *sources* em apenas um local, que permita efetuar pesquisas, recolher e devolver a informação pretendida, irá contribuir para mitigar problemas de duplicação de dados, interoperabilidade e redundância da informação.

Esta abordagem foi usada com sucesso em diferentes contextos (Fichter, 2010), utilizando diferentes conteúdos de diferentes fontes para a criação de serviços numa única interface.

Para o contexto da dissertação, um *data mashup* alimentado com a combinação dos diferentes dados dos operadores, permite a criação de uma aplicação única que devolve a informação necessária dos transportes disponíveis dos diferentes operadores, de um local a outro segmentando por janelas temporais. Assim sendo, esta abordagem irá permitir a resolução do problema de integração de diferentes esquemas de dados.

No entanto, cada fornecedor constrói os dados de forma particular, tornando a sua agregação complexa, demorada e requerendo o desenvolvimento de mecanismos de transformação de dados por programadores. Porém, investigações recentes (Hansen,

Madnick, & Siegel, 2002) (Jung, 2012) (Lee & Kim, 2012a) (Lee, 2014), apresentam composições dinâmicas através de algoritmos semânticos de aprendizagem para a criação de ontologias e agregação dinâmica.

Todavia, o processo inicia com a segmentação do pedido do utilizador (Ex: “quero ir o mais rápido possível, do local X a local Y, às 10h00 do dia de amanhã) em pequenas partes para serem respondidas pelas diferentes fontes dos operadores. De seguida, as respostas devem ser relacionadas e indexadas num repositório único, ao que irá permitir responder estruturadamente a informação dos diversos operadores.

Uma das vantagens desta abordagem, é o incremento faseado dos diferentes operadores e em diferentes escalas (regionais, nacionais, etc...) criando assim um sistema global e comum para várias redes.

2.6 Conclusão

Neste capítulo, foi abordado a atualidade dos temas em discussão e que serão abrangidos ao longo da dissertação. Foram descritos os temas, problemas que ocorrem na implementação de um *multimodal journey planner* e apresentadas soluções que os permitem dissipar.

Ficou claro que o esforço efetuado pela União Europeia na normalização dos dados não tem surtido efeito nos operadores europeus e como tal inviabiliza a adoção das regras definidas pelo CEN num contexto multimodal. A web semântica através de uma ontologia (vocabulário), permite associar os diversos operadores num único contexto de transportes públicos. Desta forma, a adição de novos operadores pode ser gradual, além de que permite a sua divulgação numa escala global (*Linked Data*). A aplicação do conceito de *API Mashup*, possibilita a integração das diferentes *API's* num único sistema sendo os termos associados a uma ontologia, que dará significado à informação recebida.

Foram discutidos os algoritmos de *routing*, sendo clarificado os problemas existentes na sua aplicação e exposto o problema com a integração de dados dos diferentes operadores, sendo sugeridos mecanismos que permitam a resolução da incerteza dos nós da rede.

Na secção seguinte é apresentada a proposta de arquitetura do *Multimodal Journey Planner* e serão apresentados os componentes que estão inseridos na prova de conceito que será desenvolvida no âmbito da dissertação.

3. *Multimodal Journey Planner*

Nesta secção, serão abordadas a visão do *multimodal journey planner*, uma proposta de arquitetura para o *multimodal journey planner* e os componentes desenvolvidos como prova de conceito. A arquitetura de um *multimodal journey planner* pode ser vista como um conjunto de vários componentes que ligados entre si, formam e compõem o sistema. Ao longo dos subcapítulos, irá ser descrito e detalhado o processo de implementação dos componentes e serão explicados os problemas que estes levantam e como eles se interligam.

A arquitetura do *multimodal journey planner* é modular, pelo que podemos caraterizar os vários componentes como módulos. Estes módulos comunicam entre si, efetuando *requests* aos serviços disponíveis. Posteriormente estes permitem a rápida integração com o MASAI, sendo que, ao longo que o MASAI vá evoluindo, poderá ser necessário ajustes para o sistema funcionar em pleno.

Numa visão de alto-nível, os componentes que constituem o *multimodal journey planner* são: 1) sistema agregador de dados; 2) cálculo da melhor rota; 3) interface móvel. O âmbito da dissertação irá abordar o principal componente, o agregador de dados e irá ser fornecido uma solução para o problema da incerteza dos nós, para que, posteriormente o cálculo da melhor rota não tenha conflitos de dados.

3.1 Visão

Como descrito no capítulo 2.1, o propósito de um *multimodal journey planner* é ajudar a obtenção de uma resposta a um utilizador de como se dirigir de um local a outro num determinado tempo. Uma aparente simples resposta a esta questão implica a obtenção de informação aos diversos operadores. Contudo, os veículos de transporte são frequentemente afetados por atrasos, acidentes, pelo que a informação estática dos horários pode afetar em muito a satisfação do utilizador/passageiro na utilização dos operadores de transporte e ainda inviabilizar o cálculo prévio da rota.

Além deste problema, devem ser considerados igualmente outros: (1) a conjugação e normalização dos diferentes dados dos operadores; (2) o conjunto dos nós de trânsito usados para concretizar o caminho desejado; (3) o número de mudanças de transporte necessárias no caminho desejado; (4) o tempo computacional para calcular a melhor solução em tempo real.

A abordagem proposta, considera a criação de um sistema que permita efetuar pedidos a diferentes operadores e conjugar a informação resultante numa única estrutura de dados semânticos sobre o qual serão efetuadas as *queries* dos utilizadores. O resultado da *query* será posteriormente processado pelo algoritmo de *routing* fornecendo depois o resultado do algoritmo à interface do utilizador.

3.2 Arquitetura

Os dados dos operadores provêm de diferentes fontes, sendo que cada operador especifica os dados de diferentes maneiras, podendo estes serem de diferentes formatos (XML, API, JSON). Como tal, é importante uma ferramenta que permita ler, interpretar e responder de forma automática os pedidos a efetuar.

O sistema proposto na sua globalidade é constituído por vários componentes, sendo o agregador um dos principais, pois irá ser o cérebro de todo o processo. O utilizador final irá comunicar com o sistema central (agregador), de forma a obter direções para se dirigir de um local a outro num determinado espaço de tempo e este, tendo em conta os diferentes operadores e após executar o algoritmo de *routing*, irá fornecer a melhor rota desde o local de origem até ao de destino numa determinada janela temporal.

Na Figura 6 é possível verificar a proposta de arquitetura para o *multimodal journey planner*. O processo inicia com o cliente a inserir as suas pretensões, para tal é necessária uma interface que permita ao utilizador inserir o local de destino e chegada, hora de partida e algumas preferências (Ex: preferência de apenas autocarro). Essa interface poderá ser uma aplicação móvel, desktop ou uma página web que servirá de ponto de ligação entre o utilizador e o sistema central.

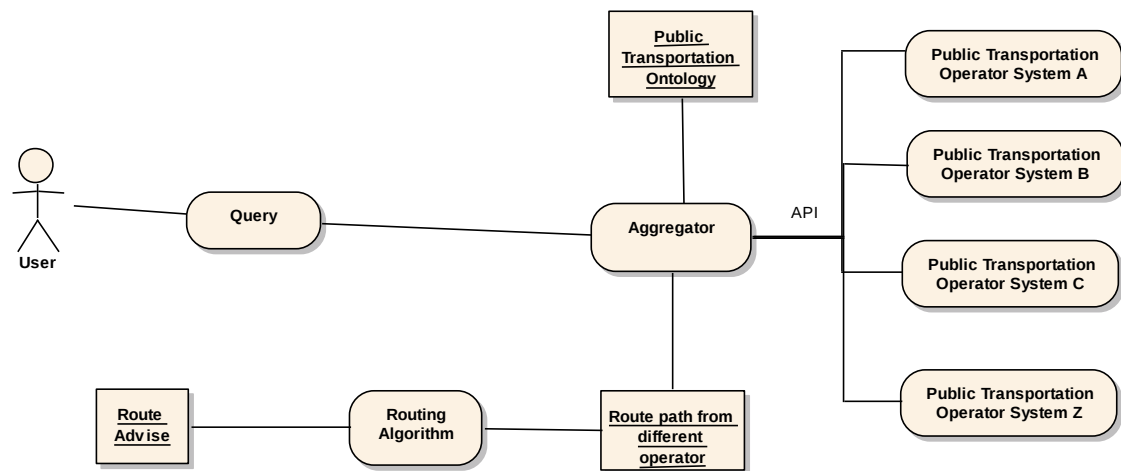


Figura 6 – Proposta da arquitetura do sistema

Após o utilizador indicar os inputs necessários, a interface enviará a informação inserida pelo utilizador para o sistema central onde este realizará uma *query* semântica às diferentes fontes de dados. A *query* previamente filtra os operadores desnecessários, permitindo assim otimizar o tempo de resposta. Todavia, como cada operador pode ter uma estrutura de dados diferentes, é necessário a existência de uma ontologia (vocabulário) que servirá como contexto a todos os transportes públicos. Desta forma, é possível garantir que os dados provenientes dos operadores respeitam a estrutura definida no sistema.

No entanto, a adaptação dos dados vindo do fornecedor é um processo bastante demorado e requer programação especializada. Estudos recentes, têm vindo a aplicar métodos semânticos em *data mashups* (Lee & Kim, 2012a) (Lee, 2014) permitindo a junção de diferentes fontes de dados, eliminando assim a necessidade de ajuste e intervenção manual para cada API.

A definição do contexto semântico, entidades e suas propriedades, relações e restrições é efetuada através de uma ontologia. A sua introdução permite além de efetuar *queries* de contexto semântico, adicionar a capacidade de promover a partilha e interoperabilidade dos dados de diferentes operadores e organizações entre os diferentes sistemas (Houda et al., 2010).

A componente principal do sistema é um agregador que é usado como componente de criação do contexto para as *queries* multioperador através das APIs disponíveis. Desta

forma a *query* é estruturada semanticamente, sendo efetuada de forma segmentada e de acordo com a estrutura da API do fornecedor de serviço. O seu output é exportado sendo depois enviado para o algoritmo de *routing* que irá apresentar a solução à interface do utilizador.

3.3 Agregador

A integração de múltiplos dados implica uma arquitetura *standard* e comum para uma comunicação e compreensão universal com as diferentes fontes de dados. Tendo isso em conta, o agregador deverá conter uma estrutura flexível com vocabulário genérico que permita ao agregador a classificação de expressões a ser utilizadas na aplicação, caracterizar as possíveis relações e definir restrições com base no uso dessas mesmas expressões. Desta forma, a mesma expressão/*query* poderá receber a informação de diferentes estruturas de dados, permitindo assim ultrapassar os problemas das diferentes estruturas de dados dos operadores.

O conceito do agregador, é considerado um *API Mashup*, onde este irá questionar as *API's* dos diferentes operadores para construir uma ferramenta que aglomera todos os dados e devolve a informação conjunta estruturada. Em suma o agregador permite: extrair informação e fornecer através de um contexto semântico um significado relacionado com as outras fontes de dados; calcular e providenciar dentro do contexto dos diferentes dados as melhores opções consoante os inputs; ferramenta de contacto entre o utilizador e os serviços dos operadores; coordenação da relação entre os dados e os serviços (Hansen et al., 2002).

A Figura 7 exemplifica a estrutura do agregador, distribuída em camadas, que ao longo do capítulo temos vindo a descrever. A primeira camada/*layer* (*Data access*) é a que servirá de ponte entre a *query* do utilizador e as fontes de informação dos fornecedores. A segunda camada/*layer* (*Analysis*) tem como propósito a extração e análise dos dados, otimizando a estrutura dos dados removendo a redundância, resolvendo conflitos e guardando numa estrutura de dados. A terceira camada/*layer* (*Transformations*) gere a transformação dos dados recebidos dos operadores para uma estrutura *standard* e contextualizada semanticamente. A quarta camada/*layer* (*Connections*) é o componente que efetuará a troca da informação entre os dados dos fornecedores e as camadas

superiores (Hansen et al., 2002). A camada dos *data sources* é representativa da informação que é recebida pelos diferentes operadores, podendo estas ser de formatos diferentes.



Figura 7 – Arquitetura do agregador

Karakostas e Kalamboukis, propõem uma API semântica de anotações onde a partir de ontologias existentes, fazem a comparação dos XSD e JSON a termos linguísticos (verbos e substantivos) mapeando assim para os conceitos ontológicos (Karakostas & Kalamboukis, 2017). O *Linked Data* fornece uma informação estruturada e descreve-a de acordo com um vocabulário específico. Através de formatos de dados específicos (RDFa⁵, MicroData⁶, JSON-LD⁷) e permitido através do *Linked Data* percorrer links/caminhos para outros pedaços de dados. Assim, é possível classificar, caracterizar e definir regras a utilizar em expressões numa aplicação.

No entanto, a mesma expressão poderá ter significados diferentes baseados no contexto corrente. Contudo, estes ao serem associadas a um *Uniform Resource Identifier* (URI) irão referenciar um termo num vocabulário específico. Com o modelo ontológico discutido do capítulo anterior, os mecanismos de descrição podem ser gerados reduzindo a replicação do mesmo conceito em múltiplos vocabulários.

⁵ <https://www.w3.org/TR/xhtml-rdfa-primer>

⁶ <https://www.w3.org/TR/microdata>

⁷ <https://json-ld.org>

3.3.1 Ontologia

Apesar dos esforços e da evolução dos formatos de standardização (discutidos no capítulo 2.2) existe uma grande diversidade de formatos, não havendo uma coesão e unificação dos mesmos termos, impossibilitando assim a sua interoperabilidade.

Num estudo europeu, Tempier e Rapp referem que a divergência dos *stakeholders* num modelo de dados e o custo de substituição dos sistemas, têm vindo a prejudicar a implementação dos *standards* numa escala internacional (Tempier & Rapp, 2011).

Como tal, podemos verificar que existe um conjunto de fatores (económicos, financeiros e empresariais) que geram relutância na aplicação de *standards* europeus na estrutura dos operadores. Do ponto de vista da empresa, uma alteração do formato de dados de um operador pode-se tornar numa operação monetariamente dispendiosa na medida que é necessário um esforço de desenvolvimento e adaptação para uma nova estrutura.

Usando o OTA⁸ (*Open Travel Alliance*) ou o PTO (*Public Transportation Ontology*) (Houda et al., 2010) são fornecidas especificações para o desenvolvimento de um modelo comum e que, além de poder fornecer contexto semântico às *APIs*, pode ainda encorajar o desenvolvimento de bibliotecas de serviços de forma a responder às expectativas dos passageiros e à indústria de transporte.

O OTA tem como missão o desenvolvimento de especificações para garantir que a informação entre o fornecedor e o turista flui nos âmbitos do turismo, hospitalidade e viagens (Open Travel, 2017). No entanto, verificou-se que a definição do modelo de dados para os transportes públicos está numa fase embrionária e de recente reestruturação, estando de momento mais vocacionada para a hotelaria e golfe.

Houda, Khemaja, Oliveira e Abed através de várias metodologias de criação de ontologias (Noy & McGuinness, 2001) (M. F. López, Gómez-Pérez, Sierra, & Sierra, 1999) (Grüninger, Fox, & Gruninger, 1995), desenvolveram e definiram o PTO (*Public*

⁸ <http://opentravel.org/>

Transportation Ontology) para representar o domínio dos transportes públicos tendo em conta a sua diversidade e os interesses dos passageiros (Houda, Khemaja, Oliveira, & Abed, 2010).

A Figura 8 representa parte da ontologia relacionada com os vários tipos de transportes na rede. Na criação da ontologia estão presentes quatro tipos de transporte (metro, comboio, eléctrico e autocarro), onde podem ser associados a várias linhas de transporte e a diferentes operadores.

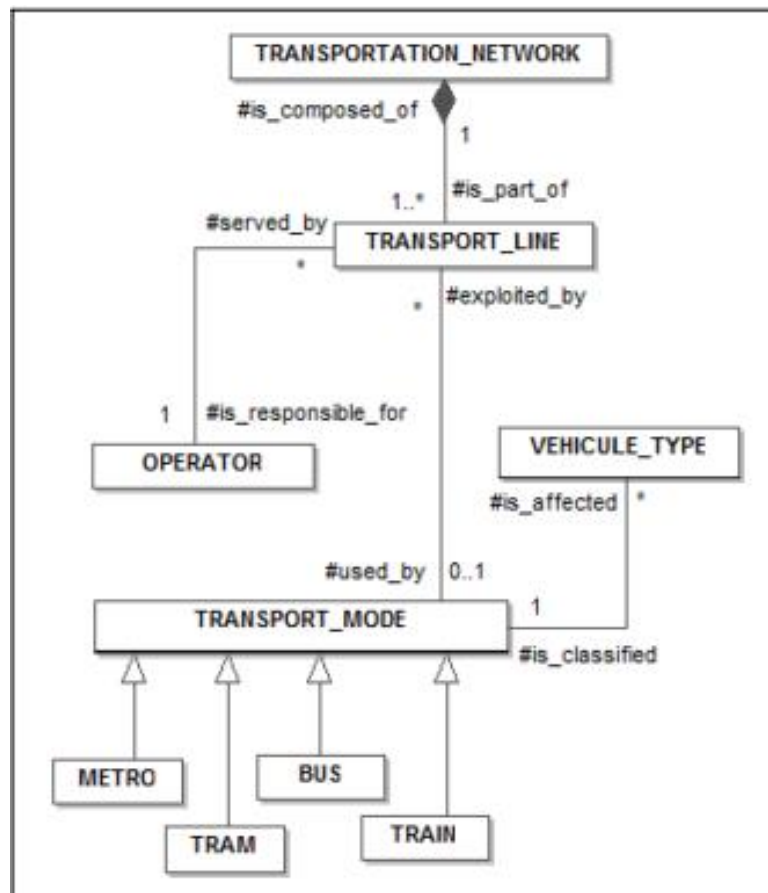


Figura 8 – Parte da ontologia relacionada com os transportes multimodais (Fonte: Houda et al., 2010)

Outro conceito instituído na ontologia e representado na Figura 9 é o *journey pattern*. Este é o conjunto ordenado dos pontos de paragem desde a origem ao destino de um percurso que associado ao *vehicle journey* define os horários de saída e chegada de cada dia da semana (Houda et al., 2010).

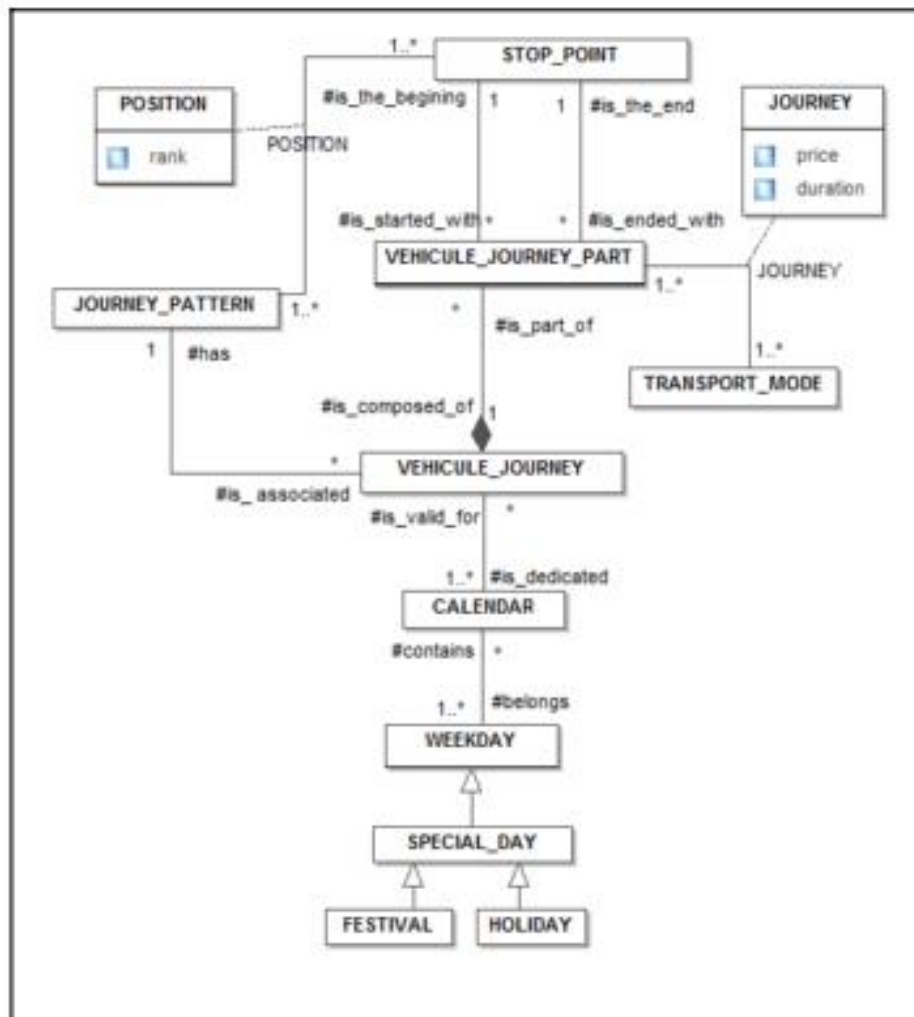


Figura 9 – Parte da ontologia relacionada com o journey planner (Fonte: Houda et al., 2010)

A definição dos principais conceitos definidos pelo PTO encontram-se descritos na Tabela 1, sendo que no anexo A, encontra-se a tabela completa definida pelo PTO.

Tabela 1 – Parte do glossário dos conceitos do PTO no original em inglês (Fonte: Houda, Khemaja, Oliveira, & Abed, 2010)

CONCEITOS	DEFINIÇÃO
CALENDAR	It allows the definition of a validity schedule period for a vehicle journey.
CONNECTION LINK	The physical (spatial) possibility for a passenger to change from one public transport vehicle to another to continue the trip.
CONNECTION POINT	A stop point where passengers change of vehicles from the same or different mode of transportation.

JOURNEY	A trip from an origin place to a destination place using a specific transport mode.
JOURNEY PATTERN	An ordered list of stop points defining one single path through the road (or rail) network.
OPERATOR	Institutions that offers the public transport.
STOP POINT	A point where passengers can board or alight from vehicles.
TRANSPORT LINE	A group of journey patterns which is generally known to the public by a similar name or number.
TRANSPORT MODE	A characterization of the operation according to the means of transport.
TRANSPORTATION NETWORK	A set of transport line to ensure public transport.
VEHICLE JOURNEY	The planned movement of a public transport vehicle on a weekday from the start point to the end point of a journey pattern on a specified infrastructure.
VEHICLE TYPE	A classification of public transport vehicles according to the vehicle scheduling requirements in mode and capacity (e.g. standard bus, double-deck).

3.3.2 Implementação

A implementação dos componentes do agregador foi segmentada por fases: 1) análise, definição e implementação da ontologia; 2) integração dos dados dos operadores na estrutura ontológica; 3) elaboração de *queries* para retornar resultados no contexto semântico; 4) resolução da incerteza dos nós da rede devido às possíveis diferentes representações dos mesmos pontos geográficos pelos operadores.

A definição da ontologia de transportes foi efetuada com o software Protégé⁹, tendo sido baseada pelos conceitos definidos por Houda et al. (2010) e Jain et al. (2014). Uma das preocupações foi a uniformização dos mesmos conceitos definidos pelos diferentes autores.

⁹ <https://protege.stanford.edu/>

O Protégé foi desenvolvido pelo *Stanford Center for Biomedical Informatics Research* da *Stanford University School of Medicine*, para permitir a criação/edição de bases de conhecimento com ontologias e a construção de modelos de domínios. Este software é *open-source* e tem sido largamente utilizado por investigadores e pela comunidade de académicos.

Na Figura 10 e Figura 11 pode-se verificar respetivamente as relações diretas definidas pelo PTO entre as classes `STOP_POINT` e `TRANSPORT_LINE` e as outras classes definidas na ontologia.

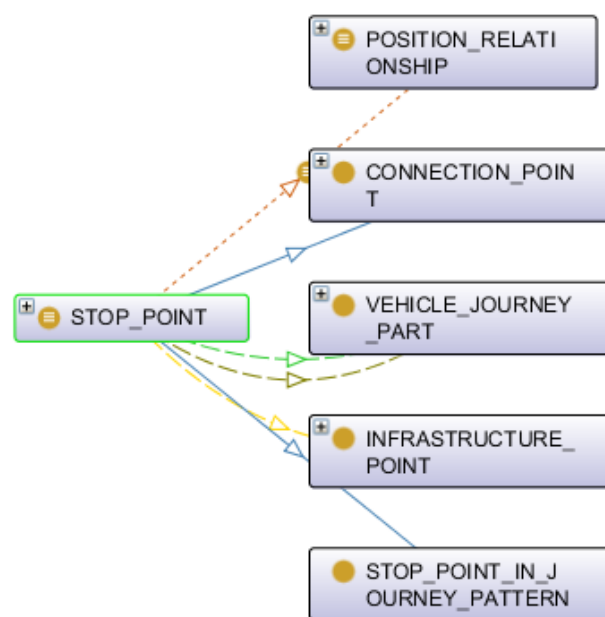


Figura 10 – Definição das classes e relações do `STOP_POINT`

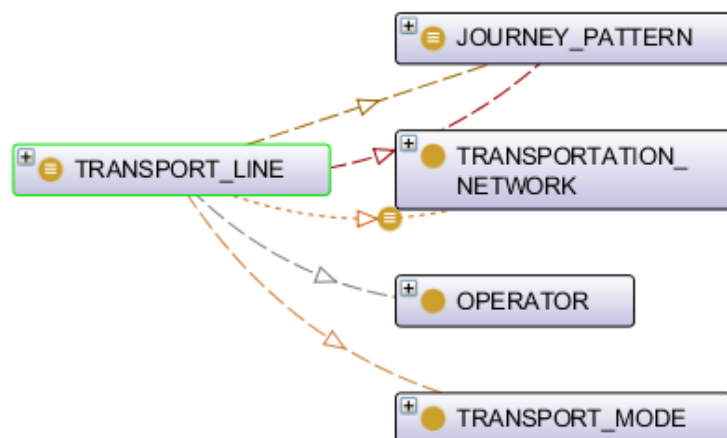


Figura 11 - Definição das classes e relações da `TRANSPORT_LINE`

Após a definição das classes, propriedades e suas relações no Protégé, é permitido a geração e exportação das classes em Java. Desta forma, estas podem ser posteriormente integradas no agregador para a criação do domínio da informação associada ao operador e ao meio de transporte.

O agregador desenvolvido como prova de conceito foi desenvolvido em *JAVA*, com uso da *framework RDF4J*¹⁰ (antigamente denominado de Sesame). O *RDF4J* permite criar, processar e guardar informação semântica no contexto do modelo. Após o *request* às *API's*, os dados recebidos dos diferentes operadores são sujeitos a um mapeamento com o objetivo de garantir que a informação X de um operador e a informação Y de um outro operador correspondam ao mesmo termo/vocabulário na estrutura de dados. Este passo é bastante importante para a interoperabilidade do sistema, pois garante a consistência e normalização da informação. Sem ele não era possível garantir o mesmo significado para diferentes conceitos vindo dos operadores.

Na Figura 12 e Figura 13 encontram-se respetivamente dois excertos em diferentes formatos da estrutura dos dados dos operadores X e Y (no apêndice A e D encontram-se os exemplos completos dos operadores). As informações de cada operador são transformadas de acordo com as limitações que algumas impõem (formatos de datas e das coordenadas) e posteriormente são guardadas no repositório de dados onde partilham a mesma estrutura de dados.

```
<?xml version="1.0" encoding="UTF-8"?>
<response>
  <Trips>
    <Trip DayType="0" StartPlaceGeo="39.4060544,-9.1343231" StartPlace="Terminal Rodoviário das Caldas da Rainha" S
    <SingleLine Arrival=" Lisboa " Depart="Caldas da Rainha " LineName="Caldas da Rainha - Lisboa via A8" IdOpera
    <Stoptimes GeoRef="39.4060544,-9.1343231" Name="Terminal Caldas da Rainha" Hour="05:30" IdLine="8247" IdStop=
    <Stoptimes GeoRef="39.3593788,-9.15743575" Name="Terminal Óbidos" Hour="05:45" IdLine="8247" IdStop="35830" I
    <Stoptimes GeoRef="39.26608922,-9.15526852" Name="Terminal Bombarral" Hour="06:00" IdLine="8247" IdStop="358
    <Stoptimes GeoRef="38.76024932,-9.15860385" Name="Terminal Campo Grande" Hour="06:45" IdLine="8247" IdStop="3
    </Trip>
    <Trip DayType="0" StartPlaceGeo="39.55028156,-8.97568658" StartPlace="Terminal Rodoviário de Alcobaça" StartTim
    <SingleLine Arrival=" Lisboa " Depart="Alcobaça " LineName="Alcobaça - Lisboa via A8" IdOperator="23" Id="825
    <Stoptimes GeoRef="39.55028156,-8.97568658" Name="Terminal Alcobaça" Hour="06:25" IdLine="8247" IdStop="35828
    <Stoptimes GeoRef="39.4060544,-9.1343231" Name="Terminal Caldas da Rainha" Hour="07:00" IdLine="8247" IdStop=
    <Stoptimes GeoRef="39.3593788,-9.15743575" Name="Terminal Óbidos" Hour="07:15" IdLine="8247" IdStop="35830" I
    <Stoptimes GeoRef="39.26608922,-9.15526852" Name="Terminal Bombarral" Hour="07:30" IdLine="8247" IdStop="358
    <Stoptimes GeoRef="38.76024932,-9.15860385" Name="Terminal Campo Grande" Hour="08:15" IdLine="8247" IdStop="3
    </Trip>
  </Trips>
</response>
```

Figura 12 – Excerto de informação recebida pelo operador X (formato XML)

¹⁰ <http://rdf4j.org/>

No apêndice B e E, encontram-se respetivamente os resultados da transformação dos dados originais dos operadores X e Y para a estrutura RDF. No apêndice F encontra-se a estrutura completa de ambos os operadores em RDF, onde se pode verificar que ambos estão inseridos num único modelo de dados.

```
[
  {
    "busNr": "1",
    "from": "Caldas da Rainha",
    "to": "Campo Grande",
    "busTrip": [{
      "unixTime": 1483257600,
      "x": 39.40590907,
      "y": -9.13474418
    },
    {
      "unixTime": 1483260900,
      "x": 38.76003844,
      "y": -9.15903009
    }
  ]
},
  {
    "busNr": "2",
    "from": "Caldas da Rainha",
    "to": "Campo Grande",
    "busTrip": [{
      "unixTime": 1483258200,
      "x": 39.40590907,
      "y": -9.13474418
    },
    {
      "unixTime": 1483261500,
      "x": 38.76003844,
      "y": -9.15903009
    }
  ]
}
]
```

Figura 13 – Excerto da informação recebida pelo operador Y (formato JSON)

No entanto, o nível da informação que os vários operadores retornam não é igual e possuem uma estrutura e formato diferentes. Torna-se necessário para casos específicos e pontuais a transformação ou rejeição da informação do operador.

Tomemos como exemplo as coordenadas geográficas das *stops*, nos casos acima demonstrado, é possível verificar que os operadores optaram pelo mesmo formato (mas podia não acontecer), no entanto é necessária uma pequena transformação para juntar a informação das coordenadas (latitude e longitude) invés de estas estarem em separado.

Outra transformação necessária, foi a transformação do formato da hora em *unixtime* para o formato de hora genérico do Java.

No mapeamento da informação recebida, para a estrutura semântica, houve a necessidade de indicar que campos dos operadores correspondem aos conceitos do PTO e em que formatos estes são guardados. Por exemplo, no caso da data/hora, o *datatype* foi atribuído como *datetime*, que irá corresponder ao *XML Schema* do W3.org.

Por fim, a informação fica registada num repositório do sistema (*SailDatabase*), permitindo ao *RDF4J* efetuar *SPARQL queries* ao banco de dados. Na Figura 14, é demonstrado um excerto da informação dos diferentes operadores em formato RDF.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns="http://aggsr.iscte-iul.pt/ontology/"
  xmlns:base="http://www.owl-ontologies.com/Ontology1256801179.owl"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:sqwr1="http://sqwr1.stanford.edu/ontologies/built-ins/3.4/sqwr1.owl#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:swrla="http://swrl.stanford.edu/ontologies/3.3/swrla.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
>

<rdf:Description rdf:about="http://aggsr.iscte-iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa"
  <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8247</ID>
  <OPERATOR_ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
  <ARRIVAL_DESTINATION rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTINATION>
  <DEPART_DESTINATION rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da Rainha</DEPART_DESTINATION>
  <START_PLACE_GEO rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544,-9.1343231</START_PLACE_GEO>
  <END_PLACE_GEO rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932,-9.15860385</END_PLACE_GEO>
  <START_TIME rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-01T05:30:00.000Z</START_TIME>
  <END_TIME rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-01T06:45:00.000Z</END_TIME>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-iul.pt/ontology/JOURNEY_PATTERN/1 1">
  <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">1</ID>
  <OPERATOR_ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">19</OPERATOR_ID>
  <ARRIVAL_DESTINATION rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Campo Grande</ARRIVAL_DESTINATION>
  <DEPART_DESTINATION rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da Rainha</DEPART_DESTINATION>
  <START_PLACE_GEO rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76003844,-9.15903009</START_PLACE_GEO>
  <END_PLACE_GEO rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.40590907,-9.13474418</END_PLACE_GEO>
  <START_TIME rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-01T08:00:00.000Z</START_TIME>
  <END_TIME rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-01T08:55:00.000Z</END_TIME>
</rdf:Description>
```

Figura 14 – Excerto da informação guardado dos operadores X e Y no formato RDF

No apêndice F encontra-se a estrutura completa do RDF, contemplando além das viagens, as *stops* que ocorrem durante o percurso do operador X até ao destino. As *SPARQL queries*, permitem a recolha dos percursos disponíveis, desde a origem até ao determinado local (facultado pelo utilizador) num determinado período de tempo. A Figura 15 fornece um exemplo de *query* para obter resultados do repositório de dados.

Os resultados do agregador serão demonstrados na secção 4, no subcapítulo seguinte será demonstrada a implementação da solução para a incerteza dos nós da rede.

```

BASE <http://www.owl-ontologies.com/Ontology1256801179.owl#>
PREFIX owl: <http://aggsr.iscte-iul.pt/ontology/>

SELECT
  ?ARRIVAL_DESTINATION
  ?DEPART_DESTINATION
  ?OPERATOR_ID
  ?ID
  ?START_PLACE_GEO
  ?START_TIME
  ?END_TIME
  ?END_PLACE_GEO
WHERE {
  ?JOURNEY_PATTERN owl:ARRIVAL_DESTINATION ?ARRIVAL_DESTINATION .
  ?JOURNEY_PATTERN owl:DEPART_DESTINATION ?DEPART_DESTINATION .
  ?JOURNEY_PATTERN owl:OPERATOR_ID ?OPERATOR_ID .
  ?JOURNEY_PATTERN owl:ID ?ID .
  ?JOURNEY_PATTERN owl:START_PLACE_GEO ?START_PLACE_GEO .
  ?JOURNEY_PATTERN owl:START_TIME ?START_TIME .
  ?JOURNEY_PATTERN owl:END_TIME ?END_TIME .
  ?JOURNEY_PATTERN owl:END_PLACE_GEO ?END_PLACE_GEO .
  FILTER (
    ?DEPART_DESTINATION='Caldas da Rainha' &&
    (?ARRIVAL_DESTINATION='Lisboa' || ?ARRIVAL_DESTINATION='Campo Grande') &&
    ?START_TIME > '2017-01-01T07:45:00.000Z'^^xsd:dateTime
  )
}

```

Figura 15 – SPARQL query para obtenção de informação dos operadores

3.4 Incerteza dos nós da rede

Como abordado no capítulo 2.4.1, a conjugação e a utilização dos dados provenientes dos diferentes operadores, gera problemas quando estes não são uniformes. A localização da mesma estação/paragem nos vários operadores pode ser diferente, levando a problemas na indicação de locais de transbordo e no cálculo algorítmico da melhor rota.

A normalização do formato geográfico foi tida em conta no capítulo anterior, no entanto, esta solução não resolve os conflitos na posição geográfica da mesma paragem/estação.

O problema é visível na Figura 16, onde várias estações/paragens são representadas por posições diferentes (Ex: ponto 1, 2 e 3). Para solucionar este problema, foi utilizado um algoritmo de *clustering* DBSCAN, que permite a criação de conjuntos de nós correspondentes a um *stop* (paragem/transbordo).

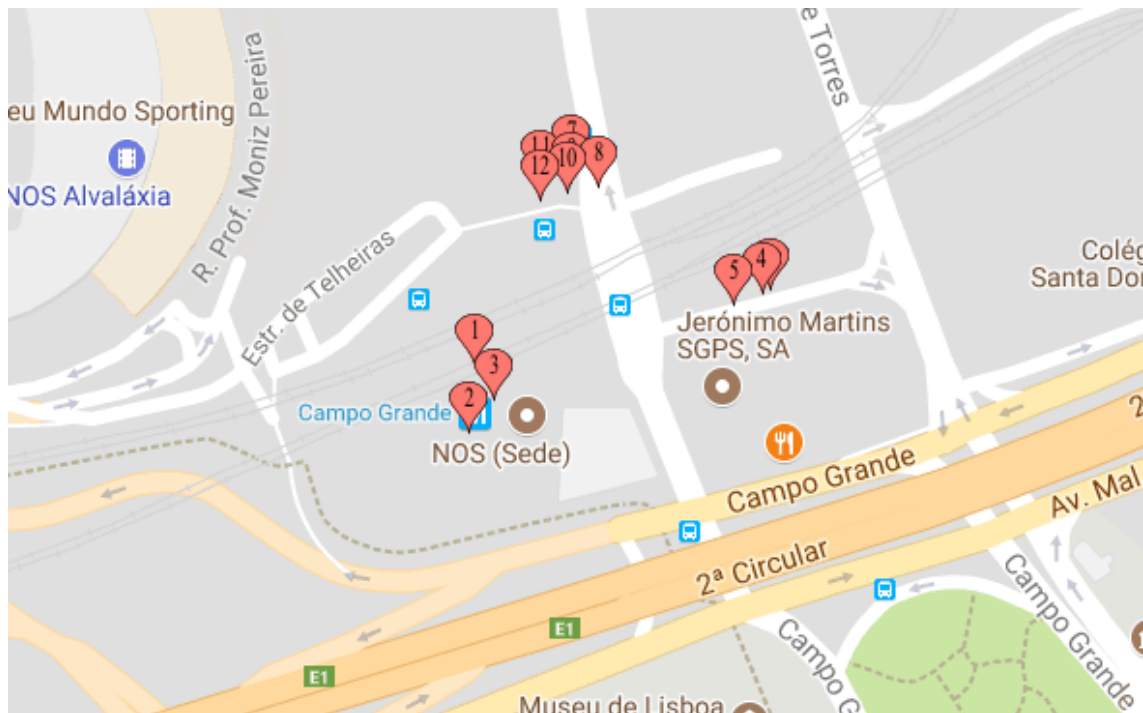


Figura 16 – Demonstração das posições fornecidas pelos operadores

O algoritmo escolhido para o *clustering* dos nós foi o DBSCAN, sendo que este apresentava mais vantagens perante o *k-means* (identificadas no capítulo 2.4.1.2). Para a sua aplicação foi utilizado o software Weka¹¹ onde foi introduzido um *dataset*, representado na Tabela 2, sendo definido como ϵ o valor 0.2 e o parâmetro *minPts* com o valor 3. Estes valores indicam que o raio entre os vizinhos é de 0.2 e o valor mínimo de vizinhos é de 3.

Após a execução do DBSCAN este identifica os vários conjuntos do *dataset*, de acordo com os parâmetros acima mencionados. Desta forma, conseguimos garantir a inclusão das diferentes paragens onde não existe coesão dos dados.

No entanto, para normalizar as paragens/estações dos vários conjuntos gerados pelo DBSCAN, foi calculado o ponto central do *cluster*. Este ponto foi calculado, efetuando uma média dos pontos desse *cluster*. Desta forma, é atribuído um ponto central (*centroid*) que é considerado como o nó principal do conjunto.

¹¹ <https://www.cs.waikato.ac.nz/ml/weka/>

Tabela 2 – Distribuição de nós ao longo de um determinado espaço

Latitude	Longitude
38.76007363	-9.15793598
38.759785	-9.15797755
38.75992095	-9.15783808
38.760379	-9.15634677
38.76033298	-9.15650234
38.76039573	-9.15629849
38.76093533	-9.15740624
38.76083913	-9.15725872
38.76086004	-9.15740624
38.7608203	-9.1574277
38.7608705	-9.15757522
38.76078266	-9.1575779

3.5 Conclusão

Os componentes/serviços foram desenvolvidos separadamente, no entanto, a sua conjugação permite a composição de parte da arquitetura do *multimodal journey planner*. O seu desenvolvimento foi efetuado com o objetivo de demonstrar que o conceito e os métodos propostos funcionam e que permitem responder aos problemas da investigação.

Em suma, ao longo deste capítulo foi apresentada e detalhada uma proposta de arquitetura de um sistema multimodal, onde foi demonstrada a implementação da prova de conceito do principal componente. Foram explicados e detalhados os problemas que se levantaram no processo de implementação, sendo sugeridas ferramentas e fornecidas soluções para os problemas em questão.

Foi explicado a aplicação e parametrização do DBSCAN para a resolução do problema de incerteza dos nós da rede, sendo que este irá trabalhar sobre a parte geográfica do *output* gerado pelo agregador.

Na secção 4 irá ser demonstrado os resultados obtidos através da implementação dos componentes/serviços descritos neste capítulo.

4. Avaliação do trabalho efetuado

Neste capítulo são expostos e avaliados os resultados obtidos pelos componentes desenvolvidos no âmbito do trabalho. A avaliação do trabalho é sobre os resultados esperados e obtidos, sendo que após a integração dos componentes no projeto MASAI é necessária uma reavaliação dos componentes.

4.1 Agregador

Como indicado anteriormente, o agregador é um dos principais componentes para o *multimodal journey planner* ao efetuar a incorporação dos vários operadores numa única estrutura/ontologia. Este deverá ter a capacidade de através de um único formato, juntar e conciliar toda a informação dispersa proveniente dos operadores, permitindo através de uma única *query* obter resultados adequados que permitam responder às questões colocadas.

Na Tabela 3 e Tabela 4 encontram-se dados de dois operadores que efetuam rotas entre as Caldas da Rainha e Lisboa (Campo Grande), sendo possível verificar que a informação original não respeita uma única estrutura de dados. Como tal, através da integração com uma ontologia, o objetivo do agregador é juntar a mesma informação numa única estrutura semântica, para posteriormente permitir através de *SPARQL queries* a extração de dados.

Tabela 3 – Parte da informação estruturada do operador X

Nome da linha	Start Place Geo	End Place Geo	Duração	Hora de Partida	Hora de Chegada	Partida	Chegada
Rapida Verde - 1	39.406(...), -9.134(...)	38.760(...), -9.158(...)	01:15	05:30	06:45	Caldas da Rainha	Lisboa – Campo Grande
Rapida Verde - 2	39.406(...), -9.134(...)	38.760(...), -9.158(...)	01:15	06:00	07:15	Caldas da Rainha	Lisboa – Campo Grande

Rapida Verde - 7	39.550(...), -8.975(...)	38.760(...), -9.158(...)	01:50	06:25	08:25	Alcobaça	Lisboa – Campo Grande
Rapida Verde - 8	39.596(...), -9.068(...)	38.760(...), -9.158(...)	02:00	06:25	08:35	Nazaré	Lisboa – Campo Grande
Rapida Verde - 9	39.406(...), -9.134(...)	38.760(...), -9.158(...)	01:15	07:20	08:35	Caldas da Rainha	Lisboa – Campo Grande

Tabela 4 – Parte da informação estruturada do operador Y

# Bus	From	To	[0] time	[0] X coord	[0] Y coord	[1] time	[1] X coord	[1] Y coord
1	Caldas da Rainha	Campo Grande	1483257600	39.405(...)	-9.13 (...)	1483260900	38.76 (...)	-9.159 (...)
2	Caldas da Rainha	Campo grande	1483258200	39.405(...)	-9.13 (...)	1483261500	38.76 (...)	-9.159 (...)
3	Caldas da Rainha	Campo grande	1483259400	39.405(...)	-9.13 (...)	1483262700	38.76 (...)	-9.159 (...)
4	Caldas da Rainha	Campo grande	1483261200	39.405(...)	-9.13 (...)	1483264500	38.76 (...)	-9.159 (...)

Na Tabela 5 é possível verificar a junção da informação dos diferentes operadores, onde esta sofreu transformação para que haja uniformidade e consistência dos dados. Neste caso, a transformação efetuada (discutido no capítulo 3.3.2) incidiu sobre o formato da data/hora e a junção das coordenadas da latitude e longitude.

Tabela 5 – Estrutura única que aglomera os dados do operador X e Y

ID	OP. ID	Dep. Dest.	Arriv. Dest.	Start Place Geo	End Place Geo	Start Time	End Time
8247	23	Caldas da Rainha	Lisboa	39.406(...), -9.134(...)	38.760(...), -9.158(...)	2017-01-01T 05:30(...)Z	2017-01-01T 06:45(...)Z
8253	23	Alcobaça	Lisboa	39.550(...), -8.975(...)	38.760(...), -9.158(...)	2017-01-01T 06:25(...)Z	2017-01-01T 08:15(...)Z
3	19	Caldas da Rainha	Campo Grande	38.760(...), -9.159(...)	39.405(...), -9.134(...)	2017-01-01T 08:30(...)Z	2017-01-01T 09:25(...)Z
4	19	Caldas da Rainha	Campo Grande	38.760(...), -9.159(...)	39.405(...), -9.134(...)	2017-01-01T 09:00(...)Z	2017-01-01T 09:55(...)Z

Na Figura 15 do capítulo anterior, encontra-se a *query* utilizada para a extração de alguns campos relevantes para a demonstração de resultados. A Tabela 6, contém os resultados obtidos através da *SPARQL query* onde são filtrados os resultados com destino a Lisboa ou Campo Grande provenientes das Caldas da Rainha a partir das 07:45.

Tabela 6 – Resultados obtidos da query ao agregador

Operador	Local de partida	Local de chegada	Hora de saída	Hora de chegada
23	Caldas da Rainha	Lisboa	2017-01-01T 08:15(...)Z	2017-01-01T 09:30(...)Z
23	Caldas da Rainha	Lisboa	2017-01-01T 09:30(...)Z	2017-01-01T 10:45(...)Z
19	Caldas da Rainha	Campo Grande	2017-01-01T 08:30(...)Z	2017-01-01T 09:25(...)Z
19	Caldas da Rainha	Campo Grande	2017-01-01T 09:00(...)Z	2017-01-01T 09:55(...)Z

Analisando os dados retornados pelo agregador, verifica-se que o resultado esperado é obtido. A *query* permitiu retornar apenas os percursos efetuados das Caldas da Rainha para o Campo Grande após as 07:45.

4.2 Incerteza dos nós da rede

Ao longo do capítulo 2.4.1, foi apresentado um problema com a incerteza dos nós da rede e no capítulo 3.4 foi apresentado uma solução para o problema em questão. Neste capítulo, será demonstrado os resultados da implementação.

Representado na Figura 16 e na Tabela 2 do capítulo 3.4, é possível verificar que os mesmos pontos (nós) são representados por posições geográficas diferentes. A execução do DBSCAN permitiu a segmentação do *dataset* (Tabela 2) em 3 diferentes *clusters*, representado na Tabela 7.

Tabela 7 – Resultado do DBSCAN na segmentação dos nós por clusters

Nós da rede	Cluster
38.760074,-9.157936	1
38.759785,-9.157978	1
38.759921,-9.157838	1
38.760379,-9.156347	2
38.760333,-9.156502	2
38.760396,-9.156298	2
38.760935,-9.157406	0
38.760839,-9.157259	0
38.76086,-9.157406	0
38.76082,-9.157428	0
38.760871,-9.157575	0
38.760783,-9.157578	0

Na Figura 17, para uma melhor demonstração dos resultados, é apresentado visualmente a distribuição dos *clusters* efetuado pelo DBSCAN.

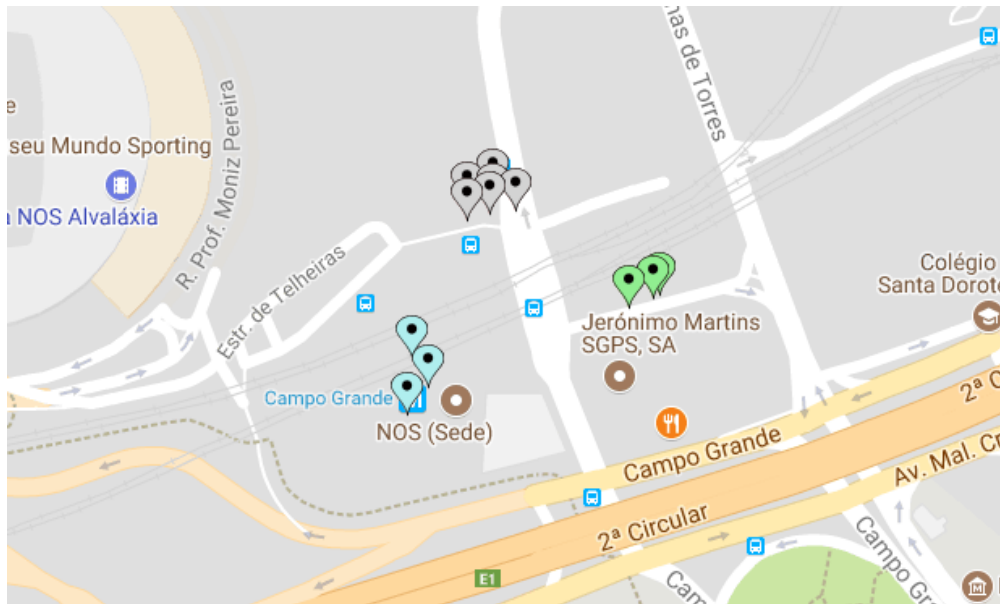


Figura 17 – Representação gráfica dos clusters criados pelo DBSCAN

Através dos *clusters* gerados e com o objetivo de definir apenas um único local geográfico (*centroid*) foi efetuado uma média das coordenadas. Este processo permitiu definir um nó chave por cada *cluster* de nós (paragens/estações). Na Figura 18 representado a vermelho encontra-se o nó principal calculado.

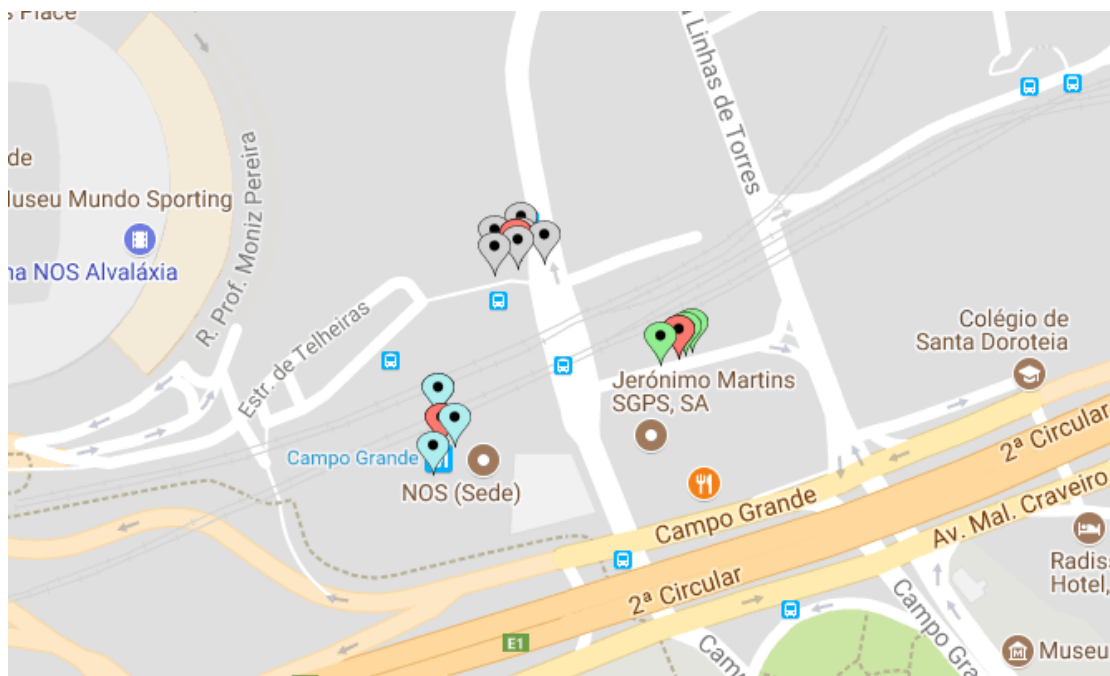


Figura 18 – Representação dos centros dos clusters (a vermelho)

4.3 Conclusão

Neste capítulo foi possível verificar que os resultados obtidos dos componentes/serviços são os esperados. O agregador permitiu a junção dos vários operadores numa única estrutura ontológica, tendo sido efetuados alguns trabalhos de transformação de dados para juntar ou adaptar dados não uniformes. Através do componente implementado, foi possível efetuar *queries* ao modelo de dados e retornar respostas dentro do contexto necessário.

A utilização do DBSCAN permitiu a junção dos vários pontos onde havia incerteza dos nós devolvidos pelos operadores, pelo que podemos afirmar que o DBSCAN lida corretamente com os pontos geográficos e que é uma opção viável para lidar com a incerteza dos nós.

Pela Figura 19 verifica-se que no âmbito da dissertação foi abordada uma parte do contexto global do MASAI. Os componentes/serviços desenvolvidos correspondem a uma parte do problema, sendo que, estes permitem posteriormente a junção de vários operadores de transportes europeus.

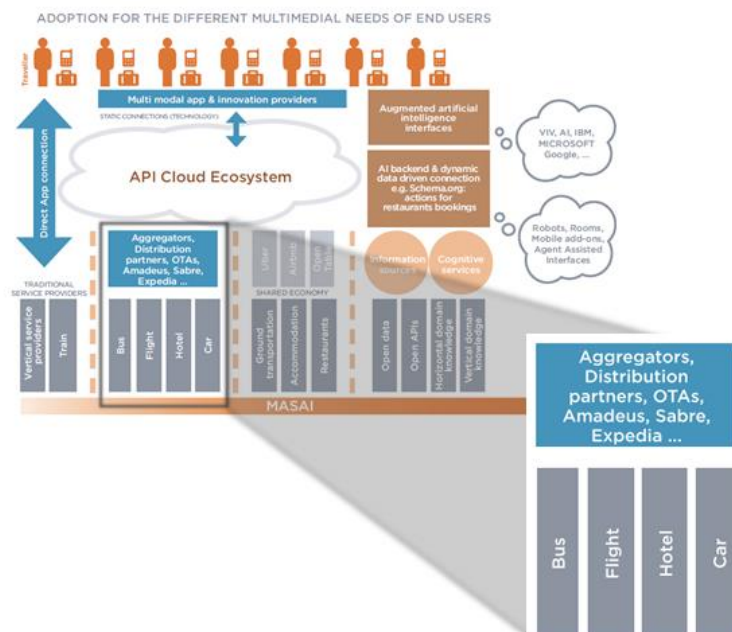


Figura 19 – Vista geral do trabalho desenvolvido no contexto do MASAI ¹²

¹² Adaptado de (ASBL, 2016)

No entanto, tendo em conta a complexidade global do projeto MASAI e que este ainda se encontra em desenvolvimento, futuramente é necessária uma reavaliação do trabalho desenvolvido com o objetivo de validar se as características se mantêm ou se a implementação necessita de ajustes, tanto ao nível da ontologia (para absorver outros tipos de operadores) como do agregador.

5. Conclusão

O desenvolvimento da dissertação teve em consideração os principais problemas e questões de investigação, tendo sido focada para os problemas que apresentam maiores dificuldades, quer por falta de investigação científica, quer por necessidade do projeto europeu.

Esta dissertação demonstra a integração de dados semânticos e a sua capacidade de resposta face às questões para obtenção de dados. Pode-se afirmar que a utilização de dados semânticos permite a integração dos dados dos operadores, contudo, esta carece de transformações para a correspondência dos termos dos operadores com os da ontologia. Ainda assim, houve dificuldade na receção dos dados dos operadores, pelo que, a integração dos dados atual não corresponde à totalidade de operadores existentes, sendo que, a sua inclusão apenas teve em conta alguns dados e é necessário que as *API's* e os dados facultados pelos operadores se estabilizem para a sua integração no agregador.

Outra limitação existente ao longo do projeto, foi que não existe uma ontologia definida pela UE. A sua discussão e definição pelos parceiros europeus irá permitir a integração universal dos dados dos operadores europeus. Contudo, é necessário ter em conta os problemas existentes com anteriores tentativas de criação de regras, guias e padrões, onde estes acabam por se tornar demasiado complexos, tornando a sua aplicação no contexto da empresa dispendioso.

O desenvolvimento da dissertação, permitiu detetar um problema na conjugação dos dados dos vários operadores, nomeadamente nas posições geográficas. Este problema foi debatido e foi apresentada uma solução que permite indicar e criar conjuntos de dados onde os nós (paragens/estações) estão compreendidos dentro de uma incerteza $\Delta\delta$.

Os componentes desenvolvidos compõem uma prova de conceito, sendo que o trabalho futuro irá permitir a integração global dos componentes, o cálculo da melhor rota e o desenvolvimento de uma interface web, desktop e móvel. Apesar dos componentes desenvolvidos serem integrados no projeto MASAI, apenas num futuro próximo teremos

feedback. Contudo, os resultados obtidos permitem constatar o potencial que existe em integrar este tipo de arquiteturas em planeadores de viagens multimodais.

Esta dissertação permitiu contribuir para a publicação do artigo “*Distributed Multimodal Journey Planner Based on Mashup of Individual Planners’ APIs*”, que foi inserido no livro (*Communications in Computer and Information Science*) publicado pela Springer. O conceito do planeador de viagens com base na integração de resultados obtidos por perguntas feitas nos diferentes operadores, foi apresentado no *Workshop DeMaDIE (Decision Making in Dynamic Information Environments)*, tendo sido incluído o agregador desenvolvido.

5.1 Trabalho Futuro

Ao longo da dissertação foram desenvolvidos alguns componentes da arquitetura proposta, no entanto, derivado à complexidade do sistema é ainda necessário trabalho futuro para concretizar a solução num todo.

O desenvolvimento da solução global, com a combinação das ligações entre os vários componentes que constituem o *multimodal journey planner*, exige a elaboração dos restantes componentes, devendo ser considerado num trabalho futuro.

O desenvolvimento da interface móvel, é uma peça importante, pois esta é a que fará a ponte entre o utilizador final, o projeto MASAI e o *multimodal journey planner*. A interface não foi elaborada no contexto desta dissertação e o seu desenvolvimento necessita de ser efetuado para existir comunicação entre o agregador e o utilizado final.

No contexto da correcção da incerteza dos nós, foram abordados os algoritmos de *routing*, sendo que é necessário implementar o algoritmo de *routing*, para que seja possível indicar ao utilizador o melhor percurso a efetuar.

A implementação do agregador, permitiu uma primeira conjugação dos dados dos operadores, no entanto, os dados inseridos não correspondem à totalidade da informação devolvida. A inserção de mais dados e de diferentes operadores irá permitir ao sistema

abranger mais transportes, uma área geográfica maior e consequentemente dar resultados mais completos e exatos aos utilizadores.

Por fim, à medida do amadurecimento do projeto MASAI, será necessária uma reavaliação do trabalho efetuado, com o objetivo de afinar algum aspeto que tenha sido alterado e também de incluir mais informação dos operadores.

Bibliografia

- Alfonso, J., Duquesne, C., Sánchez, N., Menéndez, J. M., Blanco, N., & Melero, T. (2016). Urban Mobility Data Management - The OPTICITIES Project and the Madrid Standardization Proposal. *Transportation Research Procedia*, 14, 1260–1269. <https://doi.org/10.1016/j.trpro.2016.05.198>
- Arneodo, F. (2015). *Public Transport Network Timetable Exchange (NeTEx)*.
- ASBL, M. (2016). MASAI Whitebook. Bruxelles: MASAI ASBL.
- Ben Khediri, I., Weihs, C., & Limam, M. (2012). Kernel k-means clustering based local support vector domain description fault detection of multimodal processes, 39(2), 2166–2171. <https://doi.org/10.1016/j.eswa.2011.07.045>
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. *International journal on Semantic Web and Information Systems*, 5(3), 1–22. <https://doi.org/10.4018/jswis.2009081901>
- Borole, N., Rout, D., Goel, N., Vedagiri, P., & Mathew, T. V. (2013). Multimodal Public Transit Trip Planner with Real-Time Transit Data. *Procedia - Social and Behavioral Sciences*, 104, 775–784. <https://doi.org/10.1016/j.sbspro.2013.11.172>
- Bruns, W. (2010). *SIRI (Service Interface for Real-time information)*.
- Cáceres, P., Sierra-Alonso, A., Cuesta, C. E., Vela, B., & Cervero, J. M. (2015). Modelling and linking accessibility data in the public bus network. *Journal of Universal Computer Science*, 21(6), 777–795. <https://doi.org/10.3217/jucs-021-06-0777>
- CEN/TC 278. (2008). Identification of Fixed Objects in Public Transport (TC 278 WI 00278207), 144.
- Dotsika, F. (2010). Semantic APIs: Scaling up towards the Semantic Web. *International Journal of Information Management*, 30(4), 335–342. <https://doi.org/10.1016/j.ijinfomgt.2009.12.003>
- Dudik, J. M., Kurosu, A., Coyle, J. L., & Sejdić, E. (2015). A comparative analysis of DBSCAN, K-means, and quadratic variation algorithms for automatic identification of swallows from swallowing accelerometry signals, 59 *OP-In*, 10. <https://doi.org/10.1016/j.compbiomed.2015.01.007>
- Ferreira, J. C., Andrade, A. R., Ramos, A., Martins, A. L., & Almeida, J. (2017).

- Distributed Multimodal Journey Planner Based on Mashup of Individual Planners' APIs. Em J. Bajo, Z. Vale, K. Hallenborg, A. P. Rocha, P. Mathieu, P. Pawlewski, ... J. Holmgren (Eds.), *Highlights of Practical Applications of Cyber-Physical Multi-Agent Systems: International Workshops of PAAMS 2017, Porto, Portugal, June 21-23, 2017, Proceedings* (pp. 307–314). Springer International Publishing. https://doi.org/10.1007/978-3-319-60285-1_26
- Fichter, D. (2010). What is a Mashup. *ENGARD, N. Library Mashups Exploring New Ways to Deliver Library Data*, 3–18.
- Ghiani, G., Guerriero, F., Laporte, G., & Musmanno, R. (2003). Real-time vehicle routing: Solution concepts, algorithms and parallel computing strategies. *European Journal of Operational Research*. [https://doi.org/10.1016/S0377-2217\(02\)00915-3](https://doi.org/10.1016/S0377-2217(02)00915-3)
- Gkiotsalitis, K., & Stathopoulos, A. (2015). A Mobile Application for Real-Time Multimodal Routing Under a Set of Users' Preferences. *Journal of Intelligent Transportation Systems*, 19(2), 149–166. <https://doi.org/10.1080/15472450.2013.856712>
- Grüninger, M., Fox, M. S., & Gruninger, M. (1995). Methodology for the Design and Evaluation of Ontologies. *International Joint Conference on Artificial Intelligence (IJCAI95), Workshop on Basic Ontological Issues in Knowledge Sharing*, 1–10. <https://doi.org/citeulike-article-id:1273832>
- Haghani, A., & Jung, S. (2005). A dynamic vehicle routing problem with time-dependent travel times. *Computers and Operations Research*, 32(11), 2959–2986. <https://doi.org/10.1016/j.cor.2004.04.013>
- Hansen, M., Madnick, S., & Siegel, M. (2002). *Data Integration using Web Services*. MIT Sloan Working Paper. Cambridge. https://doi.org/10.1007/3-540-36556-7_15
- Houda, M., Khemaja, M., Oliveira, K., & Abed, M. (2010). A public transportation ontology to support user travel planning. *2010 Fourth International Conference on Research Challenges in Information Science (RCIS)*, 127–136. <https://doi.org/10.1109/RCIS.2010.5507372>
- Jain, G. V., Jain, S. S., & Parida, M. (2014). Public Transport Ontology for Passenger Information Retrieval. *International Journal of Transportation Engineering*, 2(2), 131–144.

- Jung, J. J. (2012). Collaborative browsing system based on semantic mashup with open APIs. *Expert Systems with Applications*, 39(8), 6897–6902. <https://doi.org/10.1016/j.eswa.2012.01.006>
- Karakostas, B., & Kalamboukis, Z. (2017). API mashups: How well do they support the travellers' information needs? *Procedia Computer Science*, 109(2016), 204–209. <https://doi.org/10.1016/j.procs.2017.05.326>
- Lee, Y. J. (2014). Semantic-based data mashups using hierarchical clustering and pattern analysis methods. *Journal of Information Science and Engineering*, 30(5), 1601–1618.
- Lee, Y. J., & Kim, J. H. (2012a). Semantically enabled data mashups using ontology learning method for web APIs. *2012 Computing, Communications and Applications Conference, ComComAp 2012*, 304–309. <https://doi.org/10.1109/ComComAp.2012.6154862>
- Lee, Y. J., & Kim, J. S. (2012b). Automatic web API composition for semantic data mashups. *Proceedings - 4th International Conference on Computational Intelligence and Communication Networks, CICN 2012*, 953–957. <https://doi.org/10.1109/CICN.2012.56>
- López, D., & Lozano, A. (2014). Techniques in multimodal shortest path in public transport systems. *Transportation Research Procedia*, 3(July), 886–894. <https://doi.org/10.1016/j.trpro.2014.10.068>
- López, M. F., Gómez-Pérez, A., Sierra, J. P., & Sierra, A. P. (1999). Building a chemical ontology using methontology and the ontology design environment. *IEEE Intelligent Systems and Their Applications*, 14(1), 37–46. <https://doi.org/10.1109/5254.747904>
- Lozano, A., & Storchi, G. (2001). Shortest viable path algorithm in multimodal networks. *Transportation Research Part A: Policy and Practice*, 35(3), 225–241. [https://doi.org/10.1016/S0965-8564\(99\)00056-7](https://doi.org/10.1016/S0965-8564(99)00056-7)
- Noy, N. F., & McGuinness, D. L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. *Stanford Knowledge Systems Laboratory*, 25. <https://doi.org/10.1016/j.artmed.2004.01.014>
- Open Travel. (2017). About Us – Open Travel. Obtido 28 de Setembro de 2017, de <http://opentravel.org/about-us/>

- OWL Working Group. (2012). OWL - Semantic Web Standards. Obtido 1 de Setembro de 2017, de <https://www.w3.org/OWL/>
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Source Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Rocha, D. (2012). *iRouting - Utilização de técnicas inteligentes de optimização em sistemas de navegação GPS*. Faculdade de Engenharia da Universidade do Porto.
- Schubert, E., Sander, J., Ester, M., Kriegel, H.-P., & Xu, X. (2017). DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN., 42(3 OP-ACM Transactions on Database Systems. Jul2017, Vol. 42 Issue 3, p1-21. 21p.), 1. <https://doi.org/10.1145/3068335>
- Standardization, E. C. for. (2010). The World of European Standards. *Compass*.
- Standardization, E. C. for. (2013). DATEX Background | datex2. Obtido 13 de Janeiro de 2017, de <http://www.datex2.eu/content/dtex-background>
- Standardization, E. C. for. (2015). Transmodel - Reference Data Model for Public Transportation. Obtido 30 de Dezembro de 2016, de http://transmodel-cen.eu/?page_id=22
- Taheriyan, M., Knoblock, C. A., Szekely, P., & Ambite, J. L. (2014). A scalable approach to learn semantic models of structured sources. *Proceedings - 2014 IEEE International Conference on Semantic Computing, ICSC 2014*, 183–190. <https://doi.org/10.1109/ICSC.2014.13>
- Tempier, R., & Rapp, P. (2011). *Towards a European Multi-modal Journey Planner*.
- W3.org. (2007). SPARQL Query Language for RDF. Obtido 8 de Outubro de 2017, de <https://www.w3.org/TR/rdf-sparql-query/>
- Wenzel, K. (2016). Modeling with RDF and Linked Data. Obtido 21 de Setembro de 2017, de [https://www.eclipsecon.org/europe2016/sites/default/files/slides/Eclipsecon 2016 - KOMMA.pdf](https://www.eclipsecon.org/europe2016/sites/default/files/slides/Eclipsecon%202016%20-%20KOMMA.pdf)
- Ziliaskopoulos, A., & Wardell, W. (2000). Intermodal optimum path algorithm for

multimodal networks with dynamic arc travel times and switching delays. *European Journal of Operational Research*, 125(3), 486–502. [https://doi.org/10.1016/S0377-2217\(99\)00388-4](https://doi.org/10.1016/S0377-2217(99)00388-4)

Anexos e Apêndices

Anexo A – Tabela de conceitos da ontologia do *Journey Planner*

Fonte: (Houda et al., 2010)

CONCEITOS	DEFINIÇÃO
CALENDAR	It allows the definition of a validity schedule period for a vehicle journey.
CONNECTION LINK	The physical (spatial) possibility for a passenger to change from one public transport vehicle to another to continue the trip.
CONNECTION POINT	A stop point where passengers change of vehicles from the same or different mode of transportation.
EXCHANGE POLE	It is a place that aims to facilitate intermodal practices between different modes of passenger transport. The interchanges are distinguished by the variety of modes of transport they gather in one place.
GEOGRAPHIC ELEMENT	Location, place, position, site, corner, etc.
GEOGRAPHIC ELEMENT WITH SERVICE	Geographic element where different stores offer some service to the passenger or some physical structure of protection.
GEOGRAPHIC ELEMENT WITHOUT SERVICE	Geographic element where the passenger waits for a public transport.
INFRASTRUCTURE LINK	A supertype including all infrastructure points of the physical network.
JOURNEY	A trip from an origin place to a destination place using a specific transport mode.
JOURNEY PATTERN	An ordered list of stop points defining one single path through the road (or rail) network.
OPERATOR	Institutions that offers the public transport.
RAILWAY ELEMENT	A type of infrastructure link used to describe a railway network.
RAILWAY JUNCTION	A type of infrastructure point used to describe a railway network.

ROAD ELEMENT	A type of infrastructure link used to describe a road network.
ROAD JUNCTION	A type of infrastructure point used to describe a road network.
STOP POINT	A point where passengers can board or alight from vehicles.
STOP POINT IN JOURNEY PATTERN	A stop point where the passenger does not change of vehicle.
TRANSPORT LINE	A group of journey patterns which is generally known to the public by a similar name or number.
TRANSPORT MODE	A characterization of the operation according to the means of transport.
TRANSPORTATION NETWORK	A set of transport line to ensure public transport.
VEHICLE JOURNEY	The planned movement of a public transport vehicle on a weekday from the start point to the end point of a journey pattern on a specified infrastructure.
VEHICLE JOURNEY PART	A part of a vehicle journey created according to a specific functional purpose.
VEHICLE TYPE	A classification of public transport vehicles according to the vehicle scheduling requirements in mode and capacity (e.g. standard bus, double-deck).
WIRE ELEMENT	A type of infrastructure link used to describe a wire network.
WIRE JUNCTION	A type of infrastructure point used to describe a wire network.

Apêndice A – (XML) Exemplo dados do operador X

```

<?xml version="1.0" encoding="UTF-8"?>
<response>
  <Trips>
    <Trip DayType="0" StartPlaceGeo="39.4060544,-9.1343231"
      StartPlace="Terminal Rodoviário das Caldas da Rainha" StartTime="05:30"
      EndTime="06:45" Duration="01:15" Name="Rapida Verde - 1 | 05:30 Terminal
      Rodoviário das Caldas da Rainha | 06:45 Terminal | 1h15m"
      EndPlace="Terminal Campo Grande" EndPlaceGeo="38.76024932,-9.15860385">
      <SingleLine Arrival="Lisboa " Depart="Caldas da Rainha "
        LineName="Caldas da Rainha - Lisboa via A8" IdOperator="23" Id="8247"
        CompanyId="200" Version="02-05-2017 00:00:00" />
      <Stoptimes GeoRef="39.4060544,-9.1343231" Name="Terminal Caldas da
        Rainha" Hour="05:30" IdLine="8247" IdStop="35829" IsGoing="true"
        StopDestination="35830" />
      <Stoptimes GeoRef="39.3593788,-9.15743575" Name="Terminal Óbidos"
        Hour="05:45" IdLine="8247" IdStop="35830" IsGoing="true"
        StopDestination="35831" />
      <Stoptimes GeoRef="39.26608922,-9.15526852" Name="Terminal Bombarral"
        Hour="06:00" IdLine="8247" IdStop="35831" IsGoing="true"
        StopDestination="35832" />
      <Stoptimes GeoRef="38.76024932,-9.15860385" Name="Terminal Campo
        Grande" Hour="06:45" IdLine="8247" IdStop="35832" IsGoing="false"
        StopDestination="0" />
    </Trip>
    <Trip DayType="0" StartPlaceGeo="39.4060544,-9.1343231"
      StartPlace="Terminal Rodoviário das Caldas da Rainha" StartTime="06:00"
      EndTime="07:15" Duration="01:15" Name="Rapida Verde - 2 | 06:00 Terminal
      Rodoviário das Caldas da Rainha | 07:15 Terminal | 1h15m"
      EndPlace="Terminal Campo Grande" EndPlaceGeo="38.76024932,-9.15860385">
      <SingleLine Arrival="Lisboa " Depart="Caldas da Rainha "
        LineName="Caldas da Rainha - Lisboa via A8" IdOperator="23" Id="8248"
        CompanyId="200" Version="02-05-2017 00:00:00" />
      <Stoptimes GeoRef="39.4060544,-9.1343231" Name="Terminal Caldas da
        Rainha" Hour="06:00" IdLine="8247" IdStop="35829" IsGoing="true"
        StopDestination="35830" />
      <Stoptimes GeoRef="39.3593788,-9.15743575" Name="Terminal Óbidos"
        Hour="06:15" IdLine="8247" IdStop="35830" IsGoing="true"
        StopDestination="35831" />
      <Stoptimes GeoRef="39.26608922,-9.15526852" Name="Terminal Bombarral"
        Hour="06:30" IdLine="8247" IdStop="35831" IsGoing="true"
        StopDestination="35832" />
      <Stoptimes GeoRef="38.76024932,-9.15860385" Name="Terminal Campo
        Grande" Hour="07:15" IdLine="8247" IdStop="35832" IsGoing="false"
        StopDestination="0" />
    </Trip>
    <Trip DayType="0" StartPlaceGeo="39.55028156,-8.97568658"
      StartPlace="Terminal Rodoviário de Alcobaça" StartTime="06:25"
      EndTime="08:15" Duration="01:50" Name="Rapida Verde - 7 | 06:25 Terminal
      Rodoviário de Alcobaça | 08:15 Terminal | 1h50m" EndPlace="Terminal
      Campo Grande" EndPlaceGeo="38.76024932,-9.15860385">
      <SingleLine Arrival="Lisboa " Depart="Alcobaça " LineName="Alcobaça -
        Lisboa via A8" IdOperator="23" Id="8253" CompanyId="200" Version="02-05-2017
        00:00:00" />
      <Stoptimes GeoRef="39.55028156,-8.97568658" Name="Terminal Alcobaça"
        Hour="06:25" IdLine="8247" IdStop="35828" IsGoing="true"
        StopDestination="35829" />
    </Trip>
  </Trips>
</response>

```

```

    <Stoptimes GeoRef="39.4060544,-9.1343231" Name="Terminal Caldas da
Rainha" Hour="07:00" IdLine="8247" IdStop="35829" IsGoing="true"
StopDestination="35830" />
    <Stoptimes GeoRef="39.3593788,-9.15743575" Name="Terminal Óbidos"
Hour="07:15" IdLine="8247" IdStop="35830" IsGoing="true"
StopDestination="35831" />
    <Stoptimes GeoRef="39.26608922,-9.15526852" Name="Terminal Bombarral"
Hour="07:30" IdLine="8247" IdStop="35831" IsGoing="true"
StopDestination="35832" />
    <Stoptimes GeoRef="38.76024932,-9.15860385" Name="Terminal Campo
Grande" Hour="08:15" IdLine="8247" IdStop="35832" IsGoing="false"
StopDestination="0" />
  </Trip>
</Trips>
</response>

```

Apêndice B – Exemplo do RDF gerado para o operador X

```

<rdf:RDF
  xmlns="http://aggsr.iscte-iul.pt/ontology/"
  xmlns:base="http://www.owl-ontologies.com/Ontology1256801179.owl"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:sqwr1="http://sqwr1.stanford.edu/ontologies/built-
ins/3.4/sqwr1.owl#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:swrla="http://swrl.stanford.edu/ontologies/3.3/swrla.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#">

  <rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa via A8 8247">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8247</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTIN
ATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544,-
9.1343231</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932,-
9.15860385</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T05:30:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T06:45:00.000Z</END_TIME>
  </rdf:Description>

  <rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa via A8 8248">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8248</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTIN
ATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544,-
9.1343231</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932,-
9.15860385</END_PLACE_GEO>

```

```

    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T06:00:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T07:15:00.000Z</END_TIME>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa via A8 8249">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8249</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTIN
ATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544,-
9.1343231</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932,-
9.15860385</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T06:15:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T07:30:00.000Z</END_TIME>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa via A8 8250">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8250</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTIN
ATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544,-
9.1343231</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932,-
9.15860385</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T06:30:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T07:45:00.000Z</END_TIME>
</rdf:Description>

```

```

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa via A8 8258">
  <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8258</ID>
  <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
  <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTIN
ATION>
  <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
  <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544, -
9.1343231</START_PLACE_GEO>
  <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932, -
9.15860385</END_PLACE_GEO>
  <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T09:30:00.000Z</START_TIME>
  <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T10:45:00.000Z</END_TIME>
</rdf:Description>

</rdf:RDF>

```

Apêndice C – Exemplo de query para retornar as viagens do ponto Y a partir de Z horas

BASE <<http://www.owl-ontologies.com/Ontology1256801179.owl#>>

PREFIX owl: <<http://aggsr.iscte-iul.pt/ontology/>>

SELECT

?ARRIVAL_DESTINATION

?DEPART_DESTINATION

?OPERATOR_ID

?ID

?START_PLACE_GEO

?START_TIME

?END_TIME

?END_PLACE_GEO

WHERE {

?JOURNEY_PATTERN owl:ARRIVAL_DESTINATION ?ARRIVAL_DESTINATION .

?JOURNEY_PATTERN owl:DEPART_DESTINATION ?DEPART_DESTINATION .

?JOURNEY_PATTERN owl:OPERATOR_ID ?OPERATOR_ID .

?JOURNEY_PATTERN owl:ID ?ID .

?JOURNEY_PATTERN owl:START_PLACE_GEO ?START_PLACE_GEO .

?JOURNEY_PATTERN owl:START_TIME ?START_TIME .

?JOURNEY_PATTERN owl:END_TIME ?END_TIME .

?JOURNEY_PATTERN owl:END_PLACE_GEO ?END_PLACE_GEO .

FILTER (

?DEPART_DESTINATION='Caldas da Rainha' &&

(?ARRIVAL_DESTINATION='Lisboa' || ?ARRIVAL_DESTINATION='Campo Grande') &&

?START_TIME > '2017-01-01T07:45:00.000Z'^^xsd:dateTime

)

}

Apêndice D – (JSON) Exemplo dados do operador Y

```
[
  {
    "busNr": "1",
    "from": "Caldas da Rainha",
    "to": "Campo Grande",
    "busTrip": [{
      "unixTime": 1483257600,
      "x": 39.40590907,
      "y": -9.13474418
    },
    {
      "unixTime": 1483260900,
      "x": 38.76003844,
      "y": -9.15903009
    }
  ]
},
{
  "busNr": "2",
  "from": "Caldas da Rainha",
  "to": "Campo Grande",
  "busTrip": [{
    "unixTime": 1483258200,
    "x": 39.40590907,
    "y": -9.13474418
  },
  {
    "unixTime": 1483261500,
    "x": 38.76003844,
    "y": -9.15903009
  }
]
},
{
  "busNr": "3",
  "from": "Caldas da Rainha",
  "to": "Campo Grande",
  "busTrip": [{
    "unixTime": 1483259400,
    "x": 39.40590907,
    "y": -9.13474418
  },
  {
    "unixTime": 1483262700,
    "x": 38.76003844,
    "y": -9.15903009
  }
]
},
{
  "busNr": "4",
  "from": "Caldas da Rainha",
  "to": "Campo Grande",
  "busTrip": [{
    "unixTime": 1483261200,
    "x": 39.40590907,
    "y": -9.13474418
  }
]
```

```
    },  
    {  
      "unixTime": 1483264500,  
      "x": 38.76003844,  
      "y": -9.15903009  
    }  
  ]  
}
```

Apêndice E – Exemplo do RDF gerado para o operador Y

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns="http://aggsr.iscte-iul.pt/ontology/"
  xmlns:base="http://www.owl-ontologies.com/Ontology1256801179.owl"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:sqwr1="http://sqwr1.stanford.edu/ontologies/built-
ins/3.4/sqwr1.owl#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:swrla="http://swrl.stanford.edu/ontologies/3.3/swrla.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  >
  <rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/1 1">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">1</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">19</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Campo
Grande</ARRIVAL_DESTINATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76003844, -
9.15903009</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.40590907, -
9.13474418</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T08:00:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T08:55:00.000Z</END_TIME>
  </rdf:Description>

  <rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/2 2">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">2</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">19</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Campo
Grande</ARRIVAL_DESTINATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76003844, -
9.15903009</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.40590907, -
9.13474418</END_PLACE_GEO>

```

```

    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T08:10:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T09:05:00.000Z</END_TIME>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/3 3">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">3</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">19</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Campo
Grande</ARRIVAL_DESTINATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76003844, -
9.15903009</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.40590907, -
9.13474418</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T08:30:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T09:25:00.000Z</END_TIME>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/4 4">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">4</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">19</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Campo
Grande</ARRIVAL_DESTINATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76003844, -
9.15903009</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.40590907, -
9.13474418</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T09:00:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T09:55:00.000Z</END_TIME>
</rdf:Description>
</rdf:RDF>

```

Apêndice F – Exemplo do RDF gerado para o operador X e Y

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xmlns="http://aggsr.iscte-iul.pt/ontology/"
  xmlns:base="http://www.owl-ontologies.com/Ontology1256801179.owl"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:sqwrl="http://sqwrl.stanford.edu/ontologies/built-
ins/3.4/sqwrl.owl#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:swrla="http://swrl.stanford.edu/ontologies/3.3/swrla.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#">

  <rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa via A8 8247">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8247</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTIN
ATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544, -
9.1343231</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932, -
9.15860385</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T05:30:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T06:45:00.000Z</END_TIME>
  </rdf:Description>

  <rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/STOP_POINT/8247/35829">
    <ARRIVE_AT
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">05:30</ARRIVE_AT>
    <NEXT_STOP
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">35830</NEXT_STOP>
  </rdf:Description>

  <rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/STOP_POINT/8247/35830">
    <ARRIVE_AT
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">05:45</ARRIVE_AT>
    <NEXT_STOP
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">35831</NEXT_STOP>

```

```

</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/STOP_POINT/8247/35831">
  <ARRIVE_AT
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">06:00</ARRIVE_AT>
  <NEXT_STOP
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">35832</NEXT_STOP>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/STOP_POINT/8247/35832">
  <ARRIVE_AT
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">06:45</ARRIVE_AT>
  <NEXT_STOP
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">0</NEXT_STOP>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/Caldas da Rainha Lisboa via A8 8248">
  <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">8248</ID>
  <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">23</OPERATOR_ID>
  <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Lisboa</ARRIVAL_DESTIN
ATION>
  <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
  <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.4060544, -
9.1343231</START_PLACE_GEO>
  <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76024932, -
9.15860385</END_PLACE_GEO>
  <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T06:00:00.000Z</START_TIME>
  <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T07:15:00.000Z</END_TIME>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/1 1">
  <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">1</ID>
  <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">19</OPERATOR_ID>
  <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Campo
Grande</ARRIVAL_DESTINATION>
  <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
  <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76003844, -
9.15903009</START_PLACE_GEO>

```

```

    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.40590907, -
9.13474418</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T08:00:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T08:55:00.000Z</END_TIME>
</rdf:Description>

<rdf:Description rdf:about="http://aggsr.iscte-
iul.pt/ontology/JOURNEY_PATTERN/2 2">
    <ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string">2</ID>
    <OPERATOR_ID
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">19</OPERATOR_ID>
    <ARRIVAL_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Campo
Grande</ARRIVAL_DESTINATION>
    <DEPART_DESTINATION
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Caldas da
Rainha</DEPART_DESTINATION>
    <START_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">38.76003844, -
9.15903009</START_PLACE_GEO>
    <END_PLACE_GEO
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">39.40590907, -
9.13474418</END_PLACE_GEO>
    <START_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T08:10:00.000Z</START_TIME>
    <END_TIME
rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2017-01-
01T09:05:00.000Z</END_TIME>
</rdf:Description>
</rdf:RDF>

```

