



IPS Instituto
Politécnico de Setúbal
**Escola Superior de
Tecnologia de Setúbal**

**Thyago
Romagna Bendo** **DESENVOLVIMENTO DE UMA
PLATAFORMA WEB PARA A
RASTREABILIDADE DE
MODELOS BASEADOS EM
AGENTES**

Dissertação Projeto submetida como requisito
parcial para obtenção do grau de **Mestre em
Engenharia de Software**

Júri

Presidente (Doutor Cláudio Sapateiro,
ESTS/IPS)

Orientador (Especialista Cédric Grueau,
ESTS/IPS)

Vogal (Especialista José Pereira, ESTS/IPS)

Coorientador (Doutor Alan Gavioli, UTFPR)

Fevereiro, 2021

Agradecimentos

O desenvolvimento deste projeto de conclusão do curso de Mestrado de Engenharia de Software foi um período em que contei com a ajuda de diversas pessoas, as quais pela quantidade não posso citar todas, mas algumas são necessárias de tais agradecimentos:

Ao Professor Cédric Grueau, por ter orientado este trabalho, por ter dado todo o suporte necessário a execução do mesmo em tempos tão difíceis, com tão grande dedicação e disponibilidade que permitiram que este trabalho obtivesse tal resultado.

Aos meus pais, irmão e avós, por terem me incentivado em todos estes anos de vida acadêmica, me apoiando em todos os novos desafios.

Aos meus amigos, que mesmo a distância e em um momento tão difícil puderam estar ao meu lado, e aos que estiveram próximos sendo fundamentais no apoio diário.

E por fim ao Instituto Politécnico de Setúbal e a Residência de Estudantes, por terem me acolhido nesta passagem por Portugal.

Resumo

O uso de modelos baseados em agentes (MBA) é uma abordagem popular para o desenvolvimento de simulações em domínios como as ciências sociais, a geografia e as ciências naturais. Entretanto devido a complexidade dos modelos, a partilha das especificações torna-se um processo penoso, que dificulta a validação e replicação deste tipo de modelo. Para auxiliar nesse processo, foi desenvolvido o protocolo ODD que torna mais fácil a escrita e leitura das descrições dos MBA. Este protocolo evoluiu e foi refinado nas versões ODD+D e ODD+2D. Contudo, por se tratar de um protocolo, a latitude permitida na descrição dos modelos e o formato em que são disponibilizadas estas descrições dificultam, por um lado a compreensão da descrição de maneira geral, e por outro lado o próprio acesso e utilização destes modelos. As especificações dos MBA podem ser encontradas em repositórios de modelos e artigos científicos acessíveis apenas por pesquisas manuais. Para solucionar o problema foi desenvolvida uma framework, baseada numa ontologia para permitir a rastreabilidade dos modelos. Foram incluídos conceitos de outras ontologias para descrever os autores e o código informático associado a cada MBA. Foi desenvolvida uma aplicação que facilita o acesso aos modelos e que permite a análise da proveniência dos dados dos modelos e a extração das componentes do protocolo. Foi também desenvolvida uma API que permite aceder e processar as descrições de cada MBA. A aplicação foi submetida a uma série de testes que validam a abordagem usada e a utilidade da framework.

Palavras-chave: Modelos baseados em agentes, Ontologia, protocolo ODD, Aplicação Web, API, Linked Data.

Abstract

The use of agent-based models (ABM) is a popular approach to develop simulations in fields such as social sciences, geography and natural sciences.. However, due to the complexity of the models, the sharing of specifications becomes a painful process, which makes it difficult to validate and replicate this type of model. To assist in this process, the ODD protocol was developed, which makes it easier to write and read the ABM descriptions. This protocol evolved and was refined in the ODD + D and ODD + 2D versions. However, because it is a protocol, the latitude allowed in the description of the models and the format in which these descriptions are available make it difficult, on the one hand, to understand the description in general, and on the other hand, to access and use of these models. ABM specifications can be found in repositories of models and scientific articles accessible only by manual searches. To solve the problem, a framework was developed, based on an ontology to allow traceability of the models. Concepts from other ontologies were included to describe authors and document data and the computer code associated with each ABM. An application was developed to facilitate the access to the models, which allows the analysis of the source of the model data and the extraction of the protocol components. An API was also developed that allows access and process the descriptions of each MBA. The application was subjected to a series of tests that validate our approach and the utility of the framework.

Keywords: Agent-based models, Ontology, ODD protocol, Web application, API, Linked Data.

Índice

Agradecimentos	iii
Resumo	iv
Abstract	v
Índice	vi
Lista de Figuras	ix
Lista de Tabelas	x
Lista de Siglas e Acrónimos	xi
Lista de Listagens	xii
Capítulo 1	1
Introdução	1
1.1. Objetivos	2
1.1.1. <i>Objetivo geral</i>	2
1.1.2. <i>Objetivos específicos</i>	2
1.2. Justificativa	2
1.3. Organização do documento	3
Capítulo 2	4
Referencial Teórico	4
2.1. Web Semântica	4
2.1.1. <i>Abordagens RDF</i>	4
2.1.2. <i>Ontologias</i>	8
2.1.3. <i>Proveniência</i>	9
2.2. Modelos Baseados em Agentes	10
2.2.1. <i>O que são Modelos baseados em agentes</i>	10
2.2.2. <i>Aplicações de simulações</i>	11
2.3. O protocolo ODD+2D	14
2.3.1. <i>ODD</i>	14
2.3.2. <i>Revisão do ODD</i>	16
2.3.3. <i>ODD+D</i>	17
2.3.4. <i>ODD+2D</i>	18
2.4. Trabalhos Relacionados	20
2.5. Considerações Finais	20
Capítulo 3	21
Metodologia e Tecnologias	21
3.1. Metodologia de desenvolvimento	21
3.1.1. <i>Fluxograma de desenvolvimento</i>	21
3.1.2. <i>Cronograma</i>	22

3.2. Ontologia	22
3.2.1. <i>Feature Model</i>	22
3.2.2. <i>Protégé</i>	22
3.2.3. <i>Dublin Core</i>	23
3.2.4. <i>CodeMeta</i>	23
3.3. Aplicação	23
3.3.1. <i>Ferramentas</i>	23
3.3.2. <i>Estrutura</i>	24
3.4. Testes	24
3.4.1. <i>Metodologia de teste</i>	25
3.4.2. <i>Características dos testes</i>	25
3.5. Considerações finais	26
Capítulo 4	27
Resultados e Discussões	27
4.1. Ontologia	27
4.1.1. <i>Objetivo</i>	27
4.1.2. <i>Aquisição de conhecimento</i>	27
4.1.3. <i>Conceptualization</i>	29
4.1.3.1. <i>Conceitos</i>	29
4.1.3.2. <i>Relacionamentos</i>	32
4.1.3.3. <i>Atributos</i>	36
4.1.4. <i>Implementação</i>	36
4.1.5. <i>Aplicação</i>	37
4.2. Processo	38
4.2.1. <i>Visão geral do processo ODDEspec</i>	38
4.3. Avaliação	41
4.3.1. <i>Uma ferramenta de inserção, pesquisa e visualização de MBA</i>	42
4.3.2. <i>Usabilidade, Utilidade e facilidade de uso</i>	42
4.3.3. <i>Resultados e discussões</i>	42
4.4. Considerações Finais	44
Capítulo 5	45
Conclusão e Trabalhos Futuros	45
5.1. Conclusão	45
5.2. Trabalhos Futuros	46
Bibliografia	47
Apêndice I	A
Mapeamento ODD+2D parte 1	A
Apêndice II	B
Mapeamento ODD+2D parte 2	B
Apêndice III	C

Mapeamento ODD+2D parte 3C

Lista de Figuras

Figura 1 - Sublinguagens OWL.....	9
Figura 2 - Representação da emergência de padrões dos agentes [16].....	10
Figura 3 - Exemplo de mudanças no decorrer de uma simulação [1].	11
Figura 4 - Exemplo de simulação com característica espacial [16].	12
Figura 5 - Exemplo de simulação de segregação de bairros em Londres [21].	13
Figura 6 - Exemplo de análise demográfica por meio de simulação [25].	14
Figura 7 – Estrutura do protocolo ODD [26](Grimm, 2006).	16
Figura 8 - Adição do campo Data ao protocolo ODD+D [4].	18
Figura 9 – Fluxograma de execução do projeto.	21
Figura 10 – Estrutura de comunicação de dados na aplicação.	24
Figura 11 - Representação do elemento Overview na evolução de ODD ao ODD+2D no modelo de propriedades.	28
Figura 12 - Representação do elemento Design na evolução de ODD ao ODD+2D no modelo de propriedades.	28
Figura 13 - Representação do elemento Details na evolução de ODD ao ODD+2D no modelo de propriedades.	29
Figura 14 – Representação em primeiro nível da ontologia.	32
Figura 15 – Representação do elemento <i>Overview</i> na ontologia.	33
Figura 16 – Representação do elemento <i>Design Concepts</i> na ontologia.	34
Figura 17 – Representação do elemento <i>Details</i> na ontologia.	35
Figura 18 – Representação do elemento <i>Data</i> na ontologia.	35
Figura 19 – A inserção de dados de um MBA.	39
Figura 20 – Consulta de dados de um MBA.	40
Figura 21 – Acesso a API como intermédio para os MBAs.	41
Figura 22 - Taxa de sucesso nos testes de inserção e interpretação.	44

Lista de Tabelas

Tabela 1 – Cronograma de execução do projeto.	22
Tabela 2 – Os conceitos e seus atributos na ontologia.	36

Lista de Siglas e Acrónimos

API	Interface de Programação de Aplicações
HTML	Linguagem de Marcação de Hipertexto
HTTP	Protocolo de Transferência de Hipertexto
IRI	Identificador de Recursos Internacionalizado
JSON-LD	<i>JavaScript Object Notation</i> para Ligação de Dados
MBA	Modelos Baseados em Agentes
ODD	<i>Overview, Design Concepts and Details</i>
ODD+2D	<i>Overview, Design Concepts, Details, Decision and Data</i>
OWL	<i>Ontology Web Language</i>
RDF	Framework de Descrição de Recursos
RDFa	Framework de Descrição de Recursos em Atributos
REST	Transferência Representacional de Estado
URI	Identificador Uniforme de Recurso
W3C	<i>World Wide Web Consortium</i>
WHATWG	<i>Web Hypertext Application Technology Working Group</i>
XML	<i>eXtensible Markup Language</i>

Lista de Listagens

Listagem 1 - Exemplo de representação RDF/XML.	5
Listagem 2 - Exemplo de representação Trix.	5
Listagem 3 - Exemplo de representação RDFa.	6
Listagem 4 - Exemplo de representação Microdata.	6
Listagem 5 - Exemplo de representação N3.	7
Listagem 6 - Exemplo de representação Turtle.	7
Listagem 7 - Exemplo de representação N-Triples.	7
Listagem 8 - Exemplo de representação N-Quads.	7
Listagem 9 - Exemplo de representação JSON-LD.	8

Capítulo 1

Introdução

O mundo parece cada vez mais complexo e incerto. Para sustentar as suas decisões, os líderes recorrem cada vez mais a ferramentas científicas que permitem integrar estas condições de variabilidade e de complexidade. Compreender os padrões de comportamento individual, as suas causas e as consequências tem sido uma questão de relevo no domínio da Geografia nos últimos 50 anos, mas também da sociologia, da economia e de outros domínios que lidam com indivíduos. Os Modelos Baseados em Agentes (MBA) é uma das técnicas que ganhou em popularidade nas duas últimas décadas para modelar e simular a complexidade dos indivíduos. A representação feita por um modelo baseado em agentes permite que um sistema seja criado por meio de um conjunto de indivíduos autônomos capazes de tomar decisões. Essa possibilidade de modelação permite a representação de situações dinâmicas entre os indivíduos, permitindo assim a representação de diversas situações presentes no cotidiano humano [1]. Estas interações entre indivíduos são simuladas por aplicações informáticas, permitindo observar e compreender as dinâmicas entre os indivíduos e o ambiente no qual interagem.

Dentre essas possibilidades estão, por exemplo, o trabalho de [2] que aborda a simulação para o entendimento dos padrões da mobilidade humana e o trabalho de [3] que aborda o planejamento sustentável da relação entre energia, água e comida, demonstrando assim como as simulações com agentes podem ter uma eficiência grande em aspectos que muitas vezes são de difícil mensuração por outros meios.

Após duas décadas de desenvolvimento e de criação de uma comunidade alargada de utilizadores, os MBA continuam a apresentar problemas, nomeadamente no que diz respeito à comunicação. Para cada modelo desenvolvido, pode existir uma especificação diferente. As descrições de um modelo podem utilizar formalismos muito diferentes que vão da utilização da linguagem natural, até ao recurso a diagramas, passando pelas fórmulas matemáticas. O protocolo *Overview, Design Concepts and Details* (ODD) para a descrição de modelos MBA é uma primeira resposta ao problema da comunicação. Contudo, a primeira versão do protocolo está limitada pela sua capacidade em descrever apenas uma versão de modelo, a qual não possuindo uma relação direta com o código para o modelo. Esses aspectos foram aperfeiçoados com as novas versões do protocolo, com a mais recente versão *Overview, Design Concepts, Details, Decision and Data* (ODD+2D) possuindo uma capacidade descritiva consideravelmente maior [4].

Mas, apesar da melhora do protocolo descritivo, a comunicação continua a ter problemas, pois para alguém que esteja a desenvolver um MBA e deseje saber se alguma das características que tem em mente já foram desenvolvidas em outro trabalho, o modelador precisa realizar uma busca manual em uma grande quantidade de repositórios de descrições de modelos e de artigos científicos. Este estará ainda sujeito a não encontrar a informação de MBA devido a dificuldade de uma busca manual em tal escala ou a não encontrar a informação num formato adequado a interpretação.

Por isso, uma ferramenta que permita a esse desenvolvedor garantir a rastreabilidade dos MBA, de maneira a que este possa promover a divulgação, a reutilização e a validação dos seus modelos representaria um avanço para as comunidades de modeladores e utilizadores de MBA. O uso de uma abordagem ontológica para o desenvolvimento de tal ferramenta já é algo abordado em trabalhos anteriores, como o em [5]–[7] onde de diferentes maneiras, a abordagem ontológica é a escolhida para realizar a modelagem da descrição das implementações MBA. Além desses pontos já apresentados em trabalhos anteriores, a abordagem ontológica permite acrescentar uma dimensão, adicionando o conceito de proveniência dos dados, permitindo assim uma abordagem onde não somente os MBA sejam rastreáveis, como também os dados utilizados para a descrição do MBA.

Entretanto a rastreabilidade por si só não se torna relevante para que efetivamente resulte numa melhoria para o desenvolvimento MBA. Também é necessário que esta esteja inserida em uma aplicação que permita uma mais fácil relação dos usuários com os modelos.

1.1. Objetivos

1.1.1. Objetivo geral

Desenvolver uma aplicação web para a rastreabilidade de modelos baseados em agentes.

1.1.2. Objetivos específicos

- Desenvolver um modelo capaz de representar os Modelos Baseados em Agentes por meio do protocolo ODD+2D;
- Transformar este modelo em uma ontologia capaz de realizar a rastreabilidade das características dos MBA;
- Desenvolver uma aplicação que permita uma mais fácil criação e visualização de descrições de MBA do que em relação ao documento de texto;
- Desenvolver uma aplicação que seja capaz de utilizar este modelo ontológico;
- Testar a aplicação com peritos do domínio dos recursos naturais.

1.2. Justificativa

Dada a importância de análises feitas por meio de MBA, e a inexistência de uma aplicação que permita realizar a rastreabilidade de modelos de maneira a abranger todas as versões do protocolo descritivo ODD, mostra-se relevante o desenvolvimento de uma aplicação que permita aos desenvolvedores a rastreabilidade de modelos independentemente da versão do protocolo descritivo utilizada pelo modelo.

1.3. Organização do documento

Este documento está organizado da seguinte maneira: no Capítulo 2 é apresentado o referencial teórico, com os conceitos utilizados nesta tese, assim como os trabalhos que são relacionados; o Capítulo 3 irá apresentar as tecnologias que foram utilizadas e a metodologia empregada durante o trabalho; o Capítulo 4 irá apresentar os resultados e as discussões de cada um desses resultados; e por fim o Capítulo 5 irá apresentar as conclusões que foram obtidas com o trabalho.

Capítulo 2

Referencial Teórico

Este capítulo estende o capítulo introdutório, apresentando em detalhes os conceitos relacionados com o tópico desta tese. Primeiro, introduzimos os conceitos de Web Semântica e de ontologia que fornecem uma estrutura para a organização dos dados e dos conceitos de um determinado domínio. A seguir, descrevemos a abordagem do desenvolvimento baseado em modelos em engenharia de software, que foca na criação e exploração de modelos de domínio.

Vamos, também, descrever como os conceitos de Ontologia e desenvolvimento baseado em modelos se relacionam. Iremos descrever como os modelos baseados em agentes são usados pelos peritos que os implementam e introduzir o protocolo ODD, usado para documentar os modelos baseados em agentes. Vamos apresentar as características dos repositórios de modelos MBA, onde os modelos desenvolvidos são armazenados para promover a publicação e reutilização dos mesmos. Finalmente, vamos analisar como os repositórios de outros tipos de modelos científicos são geridos.

2.1. Web Semântica

A web semântica surge, de acordo com [8], como uma abordagem para trazer estrutura e significado para os conteúdos existentes na Internet, tornando esses conteúdos compreensíveis, não somente para humanos mas também para as máquinas. Para isso é preciso que o conhecimento seja adequadamente representado, de maneira que seja compreensível posteriormente. O primeiro passo para isso consiste na identificação das entidades por meio do uso do identificador uniforme de recurso (URI), para que possam ser localizadas na Internet e transmitidas por meio do Protocolo de Transferência de Hipertexto (HTTP).

2.1.1. Abordagens RDF

Esses dois elementos atuam em conjunto com um terceiro, o Framework de Descrição de Recursos (RDF), que é o responsável por representar as relações entre as diferentes URIs de maneira com que esses identificadores possam expressar algum tipo de conhecimento. O RDF tem o seu funcionamento de representação por meio do uso de triplos, compostos por um sujeito, um predicado e um objeto, com os dois primeiros itens necessariamente sendo URIs e o terceiro podendo ser tanto uma URI quanto um texto puro, que recebe a nomenclatura de literal [9].

Para garantir a correta transmissão e compreensão dos conteúdos descritos com o RDF, ele foi associado a linguagem de marcação eXtensible Markup Language (XML), que faz o uso de tags para que as entidades sejam identificadas. O uso do XML como a opção de marcação para o RDF foi um processo natural que resultou do facto de ambos serem especificações da World Wide Web Consortium (W3C), Um exemplo é apresentado na Listagem 1.

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:s="http://description.org/schema/">
  <rdf:Description about="http://www.w3.org/Home/Lassila">
    <s:Creator>Ora Lassila</s:Creator>
  </rdf:Description>
</rdf:RDF>
```

Listagem 1 - Exemplo de representação RDF/XML.

Entretanto, a abordagem RDF/XML não é a única para a representação dos metadados. Existem outras abordagens para isso. Entre elas estão: JSON-LD, Notation3, NTriples, N-Quads, Microdata, RDFa, Turtle e TriX. Cada uma delas possui características diferentes e pode ser melhor utilizada de acordo com o contexto em que são aplicadas.

Como uma evolução para a representação de grafos, a representação TriX faz o uso dos mesmos conceitos abordados na representação RDF/XML e os adapta para a que vários itens possam ser representados em conjunto, um exemplo da representação é a Listagem 2.

```
<TriX>
  <graph>
    <triple>
      <uri>http://example.org/Bob</uri>
      <uri>http://example.org/wife</uri>
      <uri>http://example.org/Mary</uri>
    </triple>
    <triple>
      <uri>http://example.org/Bob</uri>
      <uri>http://example.org/name</uri>
      <plainLiteral>Bob</plainLiteral>
    </triple>
    <triple>
      <uri>http://example.org/Mary</uri>
      <uri>http://example.org/age</uri>
      <typedLiteral
        datatype="http://www.w3.org/2001/XMLSchema#integer"
      >
        32
      </typedLiteral>
    </triple>
  </graph>
</TriX>
```

Listagem 2 - Exemplo de representação TriX.

Já com uma abordagem para a integração do conteúdo RDF com a apresentação com a

Linguagem de Marcação de Hipertexto (HTML) existe a representação Framework de Descrição de Recursos em Atributos (RDFa), permitindo a integração do conteúdo de representação visual com o de conhecimento por meio de atributos do XML, um exemplo dessa representação é apresentado na Listagem 3.

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/"
about="http://www.example.com/books/wikinomics">
  In his latest book
  <span property="dc:title">Wikinomics</span>,
  <span property="dc:creator">Don Tapscott</span>
  explains deep changes in technology,
  demographics and business.
  The book is due to be published in
  <span property="dc:date" content="2006-10-01">October 2006</span>.
</div>
```

Listagem 3 - Exemplo de representação RDFa.

Por sua vez, a representação de Microdata faz a integração da representação HTML com a representação RDF por meio de *Web Hypertext Application Technology Working Group* (WHATWG), o que tende a possuir uma maior eficiência em resultados de motores de busca, um exemplo do funcionamento dessa representação é a Listagem 4.

```
<section itemscope itemtype="http://schema.org/Person">
  Hello, my name is
  <span itemprop="name">John Doe</span>,
  I am a
  <span itemprop="jobTitle">graduate research assistant</span>
  at the
  <span itemprop="affiliation">University of Dreams</span>.
  My friends call me
  <span itemprop="additionalName">Johnny</span>.
  You can visit my homepage at
  <a href="http://www.JohnnyD.com" itemprop="url">www.JohnnyD.com</a>.
  <section itemprop="address" itemscope
itemtype="http://schema.org/PostalAddress">
    I live at
    <span itemprop="streetAddress">1234 Peach Drive</span>,
    <span itemprop="addressLocality">Warner Robins</span>,
    <span itemprop="addressRegion">Georgia</span>.
  </section>
</section>
```

Listagem 4 - Exemplo de representação Microdata.

Possuindo diversas características diferentes, a representação Notation3 ou simplesmente N3 é uma abordagem desenvolvida com o objetivo de simplificar a representação tida com RDF/XML, permitindo principalmente a redução da quantidade de caracteres utilizados e um aumento da capacidade de leitura do conteúdo representado. Isso fica evidente quando os dois formatos são comparados, como podemos ver na Listagem 5 em relação a Listagem 1:

```
@prefix dc: <http://purl.org/dc/elements/1.1/>.
<http://en.wikipedia.org/wiki/Tony_Benn>
  dc:title "Tony Benn";
```

dc:publisher "Wikipedia".

Listagem 5 - Exemplo de representação N3.

Enquanto a representação RDF/XML necessita de 279 caracteres para a representação, a abordagem N3 apresenta o mesmo conteúdo em 137. Além disso, permite a mais fácil compreensão do conteúdo por se aproximar mais da linguagem humana.

A representação N3 dá origem as representações Terse RDF Triple Language (Turtle), N-Triples e N-Quads, sendo a primeira uma simplificação da N3 com o objetivo de representar apenas conteúdo RDF, a segunda é uma simplificação para a representação de linhas que expressam o conteúdo de uma tripla (sujeito, objeto e predicado) e a terceira tem a mesma representação da N-Triple com o acréscimo de um atributo opcional para o posicionamento da tripla. Essas três representações são apresentadas nas Listagens 6, 7, e 8 respectivamente.

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
```

```
@prefix dc: <http://purl.org/dc/elements/1.1/> .
```

```
<http://www.w3.org/TR/rdf-syntax-grammar>
```

```
dc:title "RDF/XML Syntax Specification (Revised)".
```

Listagem 6 - Exemplo de representação Turtle.

```
<http://one.example/subject1> <http://one.example/predicate1>
```

```
<http://one.example/object1> .
```

Listagem 7 - Exemplo de representação N-Triples.

```
<http://one.example/subject1> <http://one.example/predicate1>
```

```
<http://one.example/object1> <http://example.org/graph3> .
```

Listagem 8 - Exemplo de representação N-Quads.

Ambas as derivações ocorrem de maneira que tudo o que é apresentado em Turtle ou em N-Triples ainda é válido como N3, mas nem todas as opções presentes na representação N3 são aceitas nas representações derivadas.

Por fim a representação de Objetos *JavaScript Object Notation* para Ligação de Dados (JSON-LD) é uma abordagem que faz uso da notação de objetos, para a representação dos dados, permitindo assim ter mais clareza em casos que há atributos dentro de outros atributos, um exemplo dessa representação é a Listagem 9.

```
{
  "@context": {
    "name": "http://xmlns.com/foaf/0.1/name",
    "homepage": {
      "@id": "http://xmlns.com/foaf/0.1/workplaceHomepage",
      "@type": "@id"
    },
    "Person": "http://xmlns.com/foaf/0.1/Person"
  },
  "@id": "https://me.example.com",
```

```
"@type": "Person",  
"name": "John Smith",  
"homepage": "https://www.example.com/"  
}
```

Listagem 9 - Exemplo de representação JSON-LD.

2.1.2. Ontologias

Entretanto, para algumas representações de conhecimento, a abordagem RDF não supre todas as necessidades. Para esses casos a abordagem por meio de ontologias é mais abrangente. As ontologias têm a sua origem no campo filosófico de análise da essência do conhecimento, permitindo assim uma abordagem mais complexa sobre um domínio [8].

Uma ontologia pode ser classificada desde altamente informal, passando por semi-informal, semi-formal até rigorosamente formal, o que acontece dependendo do contexto de uso necessário. Um dos usos onde as ontologias possuem uma grande aplicação é para unir áreas distintas que compartilham o mesmo conhecimento, permitindo assim que exista uma representação comum e portanto uma comunicação entre elas que, assim, permite garantir uma integração melhor [10].

Entretanto, o uso do modelo RDF para a definição dos conhecimentos é muito limitada, de maneira que se foi necessário o desenvolvimento uma nova tecnologia para que o conhecimento pudesse ser melhor representado nas suas diversas possibilidades. Para isso foi desenvolvida a *Ontology Web Language (OWL)*, a qual como uma linguagem de ontologia permite que os utilizadores escrevessem formalmente a conceptualização de um modelo de domínio de forma explícita [11].

A OWL possui 3 sublinguagens, as quais possuem características diferentes que devem ser escolhidas de acordo com o objetivo da aplicação. A primeira sublinguagem é a OWL Full, a qual é totalmente compatível com o formato RDF, mas tão poderosa que é indecível; já a segunda sublinguagem é a OWL DL, que abre mão de parte da compatibilidade com o formato RDF para ser mais racionalizáveis; por fim se tem a sublinguagem OWL Lite, que abre mão de ainda mais elementos da compatibilidade com o formato RDF, mas se torna muito mais acessível a quem a utiliza [12]. A compatibilidade entre as diferentes sublinguagens pode ser vista melhor no diagrama apresentado a Figura 1:

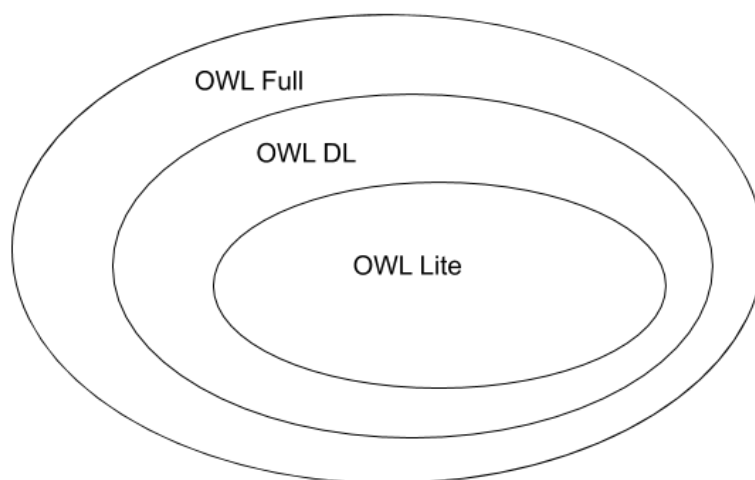


Figura 1 - Sublinguagens OWL.

Entre as possibilidades existentes nas ontologias está a integração de schemas externos para abordar uma representação, permitindo assim que um conteúdo já testado e avaliado seja agregado na representação nova. Essa alternativa além de agregar a qualidade que é derivada do schema, permite que somente o que seja interessante para a ontologia nova possa ser utilizado, evitando a representação desnecessária de certos itens na especificação.

2.1.3. Proveniência

Proveniência é a capacidade de registrar o histórico dos dados e o seu local de origem. Em geral, é a capacidade de determinar a cronologia da propriedade, custódia ou localização de qualquer objeto.

No contexto da proveniência dos dados, a proveniência documenta as entradas, entidades, sistemas e processos que influenciam os dados de interesse, fornecendo de fato um registro histórico dos dados e das suas origens.

A PROV-O é um dos exemplos das integrações possíveis. Trata-se de uma das ontologias desenvolvidas pelo World Wide Web Consortium. É uma ontologia dedicada a proveniência dos dados, a qual pode ser definida como o registo que descreve as pessoas, instituições, entidades e atividades com participação em produzir, influenciar ou entregar os dados [13].

A PROV-O se tornou um tema de estudo de interesse devido à necessidade de um paradigma para simplificar a implementação da proveniência dos dados, sendo essa uma necessidade de busca em pesquisas por 3 décadas. PROV-O é uma abordagem que conseguiu atingir a boa integração aos sistemas e extensibilidade [13].

2.2. Modelos Baseados em Agentes

2.2.1. O que são Modelos baseados em agentes

[14] descreve a simulação como um processo importante, pois é por meio dela que se torna possível modelar, situações que possuem uma complexidade que não permite criar modelos somente por meio de uma abordagem analítica.

Com o desenvolvimento computacional e uma maior quantidade de dados disponíveis, foi possível com que as abordagens para simulação evoluíssem, essa evolução levou ao desenvolvimento e crescimento da simulação por meio de agentes, sendo a abordagem com maior desenvolvimento os agentes baseados em modelos [15].

A definição de modelos baseados em agentes é dada por [16] como uma simulação de computador na qual os indivíduos são representados por agentes heterogêneos de diferentes tipos capazes de tomar de decisões autonomamente. Estes agentes obedecem a regras para se relacionar com outros agentes e o ambiente que lhes é fornecido, permitindo a emergência de padrões, como apresentado na Figura 2.

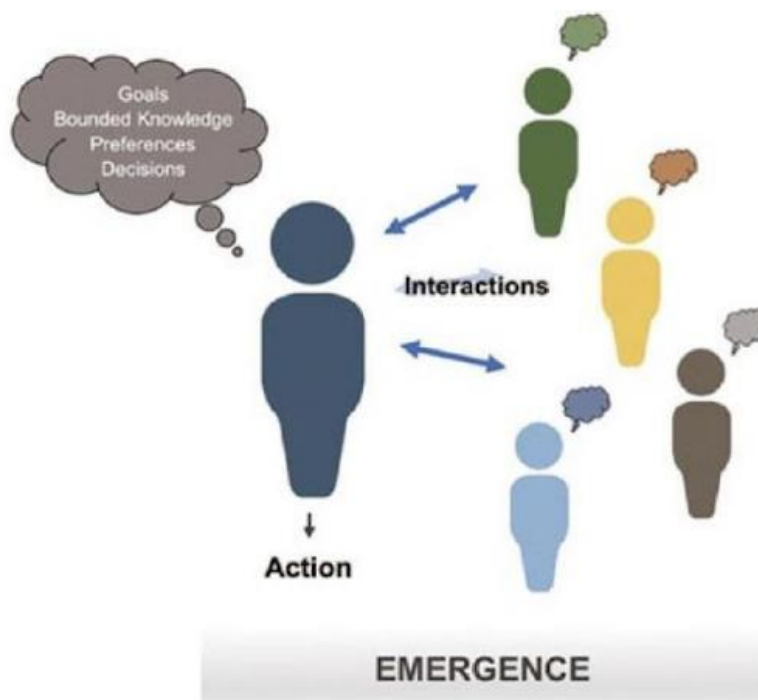


Figura 2 - Representação da emergência de padrões dos agentes [16].

Para a definição de um agente, alguns itens são fundamentais, esses itens são: os agentes são autônomos, agem necessariamente fora de um controle central; os agentes podem possuir a heterogeneidade, para permitir que a simulação apresente uma dinâmica entre grupos diferentes; os agentes tomam decisões independentemente, podendo isso ser de acordo com um objetivo, de forma reativa, racionalmente, para comunicação ou mobilidade; e agentes podem possuir a habilidade de aprender e se adaptar no decorrer da simulação [17].

As regras aplicadas aos agentes são os elementos responsáveis por guiar o seu

comportamento e decisões, sendo das regras a responsabilidade direta pela dinâmica que a simulação tende a tomar. As regras geralmente são formuladas no formato se-então-senão, com as decisões sendo tomadas após a satisfação de uma das condições, essas condições podem ser formuladas de maneira a dar ao agente o conceito de aprendizagem [15].

As simulações por meio de agentes são descritas como vantajosas por [1] por permitirem a descrição da descontinuidade do comportamento do indivíduo, o que se dá devido à característica não linear e descontínua da interação dos agentes; a heterogeneidade dos agentes e das suas interações permitem a representação de populações distintas em conceitos e racionalidade e de efeitos em rede de acordo com as diferentes interações; e, pois, devido à capacidade de aprender e se adaptar, agentes podem representar diversas situações de grande complexidade, como, por exemplo, a representação de uma segregação entre indivíduos, na Figura 3.

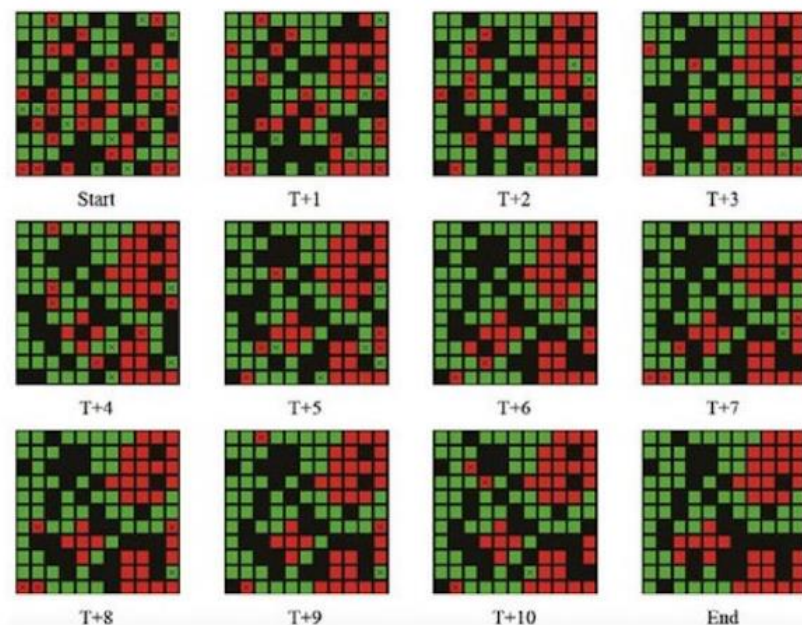


Figura 3 - Exemplo de mudanças no decorrer de uma simulação [1].

Os MBA podem ser aplicados em diversas áreas distintas, como para a modelação do comportamento de uma bolsa de valores, a predição da disseminação de epidemias, a compreensão do comportamento de consumidores, o comportamento de uma civilização antiga na sua queda ou até mesmo o comportamento num campo de batalha. A modelação de um agente pode ser extremamente simples, considerando somente aspetos fundamentais para a simulação, ou extremamente complexa, considerando os detalhes mínimos para que os agentes tenham maior profundidade nas suas decisões [18].

2.2.2. Aplicações de simulações

O uso de simulações com MBA nos domínios relacionados com à Informação Geográfica, a Modelação Espacial e a Simulação Espacial são processos conduzidos por utilizadores de

um sistema de informação, um modelador, a fim de entender fenômenos espaciais e planejar com antecedência.

Entre as várias possibilidades de uso, podemos referir a exploração das alternativas presentes num cenário, para a compreensão do funcionamento de determinada dinâmica, como também a integração de um modelo de simulações num sistema mais amplo de apoio à decisão espacial.

Um dos primeiros modelos a demonstrar a capacidade de aplicação de MBA e o aparecimento de informações desses modelos foi o SugarScape de [19] que apresentou a aplicação de uma simulação social com presença espacial, este modelo continua a ser explorado, como por [16], apresentado na Figura 4.

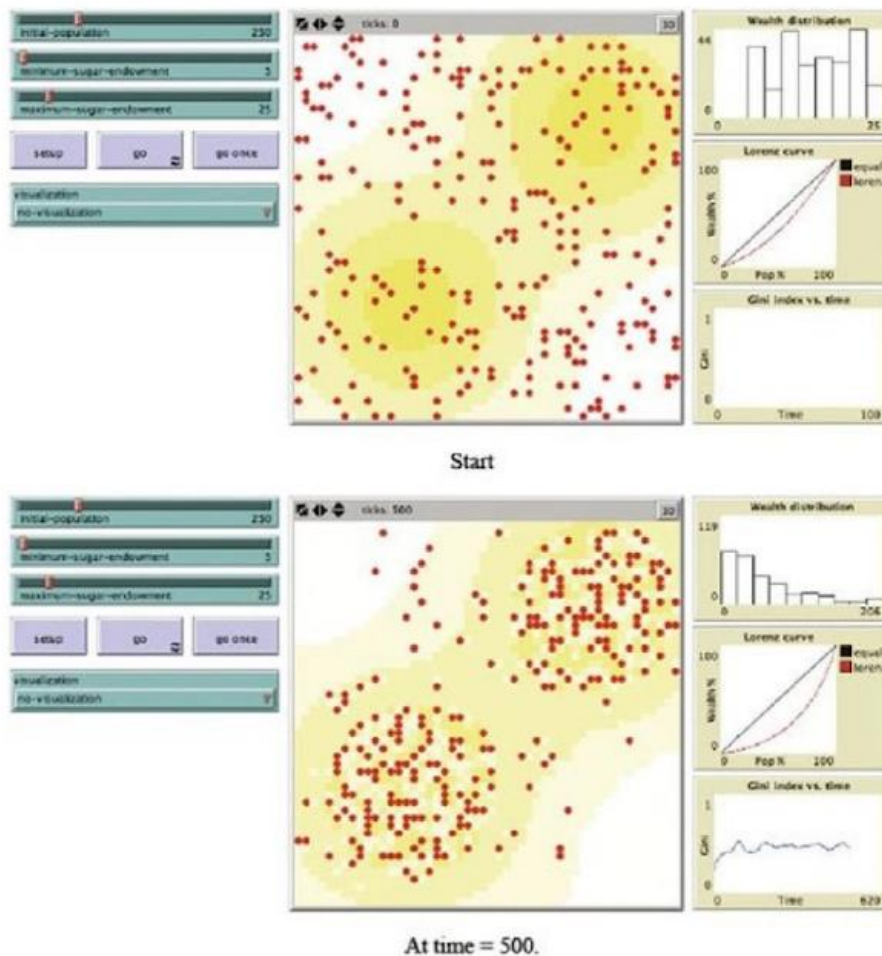


Figura 4 - Exemplo de simulação com característica espacial [16].

Essas possibilidades fazem da simulação um elemento-chave para o estudo da mudança no uso da terra e para o apoio a tomada de decisões sobre sistemas socioeconômicos e ecológicos. Os modelos baseados em agentes permitem que os modeladores espaciais representem as atribuições, decisões e ações de vários atores que afetam paisagens espaciais, podendo assim melhor entender a dinâmica presente entre os diferentes atores.

Por meio da simulação, os especialistas podem reproduzir as interações dos agentes ao longo do tempo para definir rotas de evacuação quando ocorrer um desastre, como um terremoto [20], ou para prever mudanças no uso da terra, como, por exemplo para prever a

expansão urbana de uma determinada área em 20 anos, ou lidar com restrições da população no acesso à água de determinada bacia hidrográfica.

A simulação pode servir como um revisor de teoria, por exemplo, para verificar uma teoria de como a segregação ocorre nos bairros da cidade, por exemplo [21] com um estudo sobre Londres e as diferenças entre regiões e microrregiões apresentado na Figura 5. Nas últimas décadas, o grau de flexibilidade e a capacidade de representar interações sociais induziram equipes multidisciplinares que trabalham em ciências geográficas a desenvolver diferentes modelos baseados em agentes. [22] e [23] fornecem uma ampla variedade desses aplicativos, que vão da exploração da dinâmica urbana na América Latina.

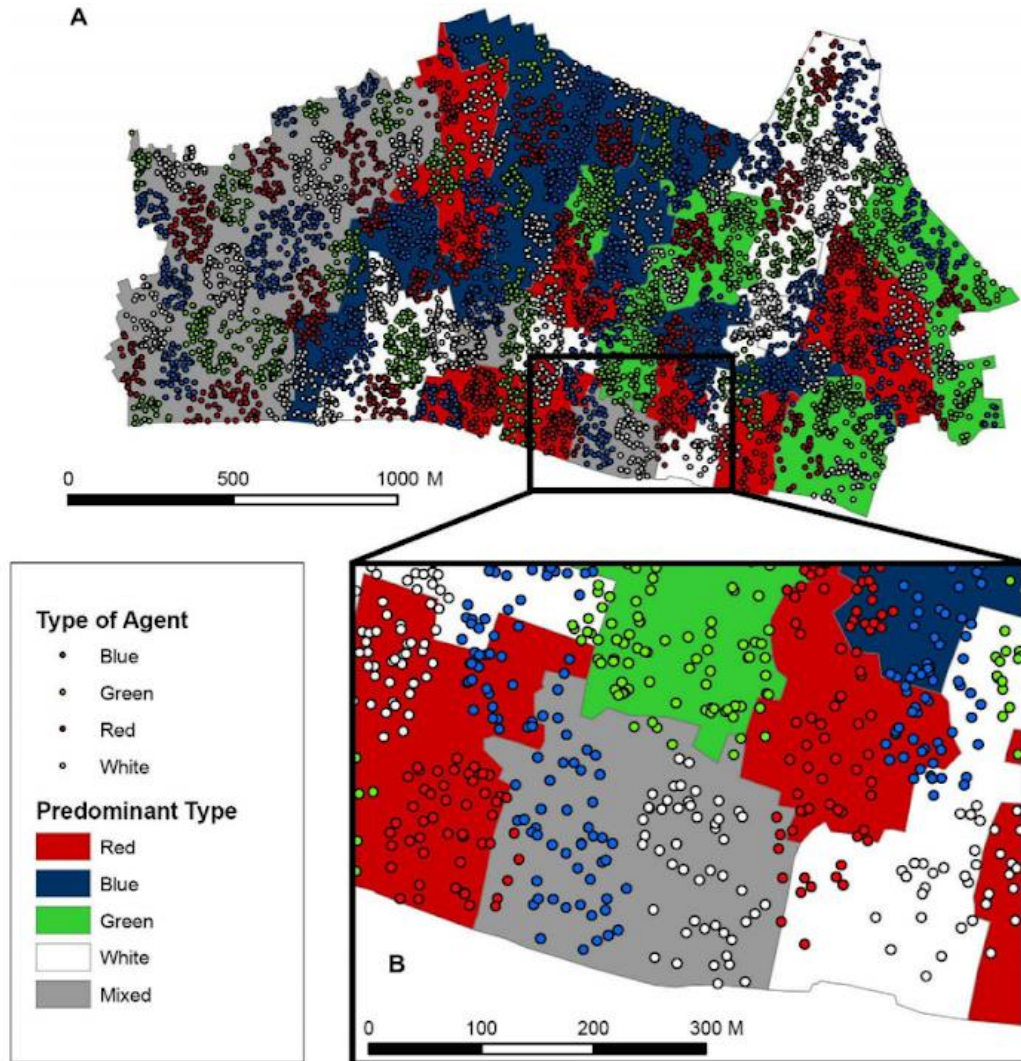


Figura 5 - Exemplo de simulação de segregação de bairros em Londres [21].

Outros estudos de da dinâmica urbana pode ser traçado desde simulações em relação ao tráfego de Chicago em 1955 [22], e atualmente se relaciona com a propagação espacial de epidemias, o entendimento da mobilidade humana [2](González, 2008), ou a análise do comportamento de veículos elétricos e a necessidade de recarga numa cidade [24].

Mas as análises vão além de possibilidades urbanas, podendo incluir a gestão de florestas, ou as dinâmicas da tecnologia agrícola, já que as ciências sociais, demográficas

(Figura 6), económicas e ambientais também domínios são de aplicação para os MBA [25].

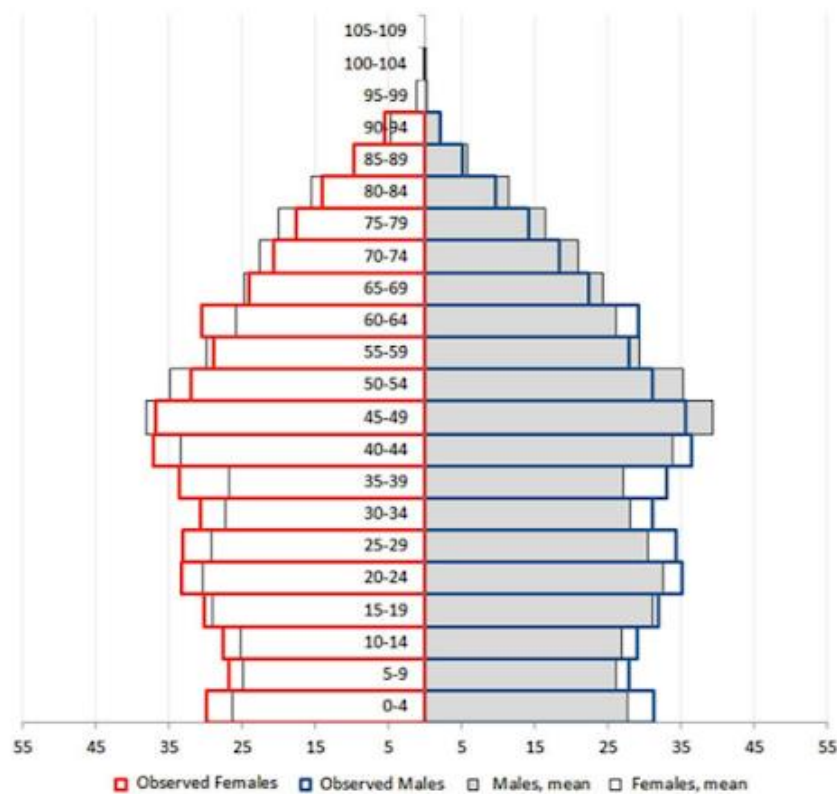


Figura 6 - Exemplo de análise demográfica por meio de simulação [25].

2.3. O protocolo ODD+2D

Existem várias abordagens para se descrever os dados da modelação dos agentes. Para tentar unificar a descrição, foi proposto o protocolo *Overview, Design Concepts and Details*. Este protocolo foi estendido com a adição do campo *Decision* para formar o protocolo ODD+D, por fim o protocolo foi aperfeiçoado com a adição do campo *Data* referido pelo protocolo ODD+2D. Na sequência será abordada detalhadamente todos os processos dessa evolução.

2.3.1. ODD

O protocolo ODD foi proposto por Grimm e um vasto conjunto de colegas [26] com o propósito de unificar a descrição de modelos baseados agentes, de maneira a permitir que a compreensão da abordagem utilizada na modelação de um determinado sistema de agentes seja facilitada. Para alcançar isso os autores utilizam uma divisão básica: *Overview, Design Concepts, e Details*, que dá origem ao nome do protocolo.

O elemento de *Overview* é destinado a descrição inicial dos agentes, apresentando os conceitos básicos e os objetivos do modelo. O seu conteúdo é dividido em 3 campos: *Purpose; State variables, and scales; e Process Overview and Scheduling*.

O elemento de *Purpose* é usado para a apresentação da proposta do trabalho, justificando

as escolhas de abordagem para a formulação do problema, para que quem esteja lendo o trabalho possa entender o porquê de alguns aspetos serem considerados e outros não.

Já o elemento *State variables, and scales* é responsável por apresentar a estrutura do modelo em construção, apresentando todas as variáveis presentes no modelo, conjuntamente com as características e os objetivos de cada variável, além de apresentar como é a mensuração no modelo.

Por fim o elemento de *Process Overview and Scheduling* é responsável por conter uma descrição de todas as etapas presentes no processo, com os efeitos que o mesmo apresenta no restante do modelo, além da apresentação de como se dá a relação entre os processos e as alterações que cada um possui nas variáveis.

Na sequência o elemento *Design Concepts* é responsável pela descrição das abordagens de implementação escolhidas para o projeto. O seu conteúdo é composto somente por um elemento, mas que possui diversos itens onde os conteúdos são distribuídos. Esses itens são: *Emergence; Adaptation; Fitness; Prediction; Sensing; Interaction; Stochasticity; Collectives; e Observation.*

O item de *Emergence* descreve que fenômenos emergem das interações individuais dos elementos e quais são impostos. Já o item de *Adaptation* possui a descrição da forma como os indivíduos do modelo são capazes de se adaptar e melhorar a sua aptidão. Na sequência o item *Fitness* define se a avaliação é implícita ou explícita, descrevendo a maneira como deve ser medida no caso de explícita.

Por sua vez o item *Prediction* apresenta como os indivíduos avaliam as consequências dos seus atos e experiências que presenciaram. Já o item *Sensing* apresenta quais variáveis os indivíduos utilizam para as suas decisões de adaptação. Na sequência o item *Interaction* apresenta que interações entre os indivíduos são aceitas.

Já o item *Stochasticity* descreve a componente estocástica do modelo e a justificativa da sua utilização. Por sua vez o item *Collectives* apresenta se os indivíduos de alguma maneira são agrupados num coletivo. Por fim o item *Observation* apresenta como dos dados são coletados dos agentes para que eles sejam testados, entendidos e analisados.

Por fim o elemento de *Details* é responsável pela apresentação dos detalhes técnicos que foram utilizados na abordagem de implementação. O conteúdo é dividido em 3 elementos, que são: *Initialization; Input; e Submodels.*

O elemento *Initialization* é responsável por indicar o estado inicial do modelo, como, por exemplo como é inicializado o ambiente para a simulação, quais os valores das variáveis, se a inicialização é baseada em dados ou não e se a inicialização varia com o decorrer das simulações.

Por sua vez o elemento de *Input* que apresenta quais são os dados utilizados em entrada para o modelo, para que o modelo apresente uma maior capacidade dinâmica e se aproxime da situação real que deseja-se simular.

Por fim o item de *Submodels* contém a demonstração de como os submodelos contidos no funcionamento do modelo funcionam, com a parametrização utilizada para que os

processos decorram, com a descrição ocorrendo de acordo com as possibilidades para o caso do modelo.

Overview	Purpose
	State variables and scales
	Process overview and scheduling
Design concepts	Design concepts
Details	Initialization
	Input
	Submodels

Figura 7 – Estrutura do protocolo ODD [26](Grimm, 2006).

2.3.2. Revisão do ODD

Os autores do protocolo ODD realizaram uma análise da utilização do protocolo [27], avaliando como a utilização feita pelos utilizadores passados uns anos. Além desta análise, apresentaram alterações ao protocolo para melhor a adequação às necessidades dos utilizadores. As alterações foram as seguintes:

Mudança de *State variables and scales* para *Entities, state variables, and scales*, adicionando a descrição de todas as entidades presentes no modelo, sendo uma entidade definida como um objeto ou ator distinto que pode interagir com outras entidades, ou ser afetado pelo ambiente.

Adição do item *Basic Principles* para o elemento *Design Concepts*, com o objetivo de realizar uma introdução dos conceitos gerais da abordagem escolhida e das hipóteses que estão sob avaliação.

Mudança do item *Fitness* para *Objectives* no elemento de *Design Concepts*, para que fique mais claro o conteúdo que deve ser contido neste item, pois apesar de as palavras serem consideradas sinônimos em diversos contextos, para a descrição de agentes *Objectives* contempla mais possibilidades.

Adição do item *Learning* ao elemento de *Design Concepts*, para que seja descrita a forma como os agentes fazem a sua aprendizagem, considerando inclusive possíveis mudanças na aprendizagem devido a qualquer fator que seja.

Mudança do elemento de *Input* para *Input Data*, com o objectivo de representar melhor o conteúdo do que deve ser apresentada no elemento, pois os dados em si são também parte importante no elemento.

Com essa atualização, o protocolo apresentava o seguinte formato:

- *Overview;*
 - *Purpose;*
 - *Entities, state variables, and scales;*
 - *Process Overview and Scheduling;*
- *Design Concepts:*
 - *Basic Principles;*
 - *Emergence;*
 - *Adaptation;*
 - *Objectives;*
 - *Learning;*
 - *Prediction;*
 - *Sensing;*
 - *Interaction;*
 - *Stochasticity;*
 - *Collectives;*
 - *Observation;*
- *Details*
 - *Initialization;*
 - *Input Data;*
 - *Submodels.*

2.3.3. ODD+D

Já [28] apresenta uma evolução do protocolo, de ODD para ODD+D, com a adição do campo de *Decision*, com as seguintes alterações:

Mudança do item de *Basic Principles* para *Theoretical and Empirical Background* ao elemento de *Design Concepts*, de forma a deixar mais claro como o conteúdo do elemento deve se referir às questões em que os agentes se baseiam. Os conceitos de *Objectives* e *Adaption* foram transferidos para a criação do conceito de *Individual Decision Making* no elemento de *Design Concepts*, de maneira a demonstrar uma maior abrangência do conteúdo presente no elemento, podendo este conter toda a descrição do processo de decisão dos indivíduos.

Adição do item de *Heterogeneity* ao elemento de *Design Concepts*, para que em caso de haver uma heterogeneidade entre os agentes, a mesma possa ser descrita e que seja apresentado o que difere os agentes. Mudança dos itens *Prediction* e *Sensing* para *Individual Prediction* e *Individual Sensing* respetivamente, ambos no elemento de *Design Concepts*, para que a nomenclatura evidencie que o conteúdo se refere às questões individuais. O conceito de *Observation* foi expandido, com a integração do conceito de *Emergence*, pois os conceitos de ambos foram considerados próximos, e os autores avaliaram que a separação causava um

maior risco de má compreensão do que um benefício descritivo. Adição do elemento de *Implementation Details* como subelemento de *Details*, para que os detalhes da implementação possam ser apresentados adequadamente, se possível fornecendo acesso a implementação.

Após as alterações, o protocolo se encontrava da seguinte maneira:

- *Overview*;
 - *Purpose*;
 - *Entities, state variables, and scales*;
 - *Process Overview and Scheduling*;
- *Design Concepts*:
 - *Theoretical and Empirical Background*;
 - *Individual Decision Making*;
 - *Learning*;
 - *Individual Sensing*;
 - *Individual Prediction*;
 - *Interaction*;
 - *Collectives*;
 - *Heterogeneity*;
 - *Stochasticity*;
 - *Observation*;
- *Details*
 - *Implementation Details*;
 - *Initialization*;
 - *Input Data*;
 - *Submodels*.

2.3.4. ODD+2D

Por fim [4] apresenta a mais recente evolução do protocolo, de ODD+D para ODD+2D, adicionando o campo de *Data*, como apresentado na Figura 8, com as seguintes alterações:

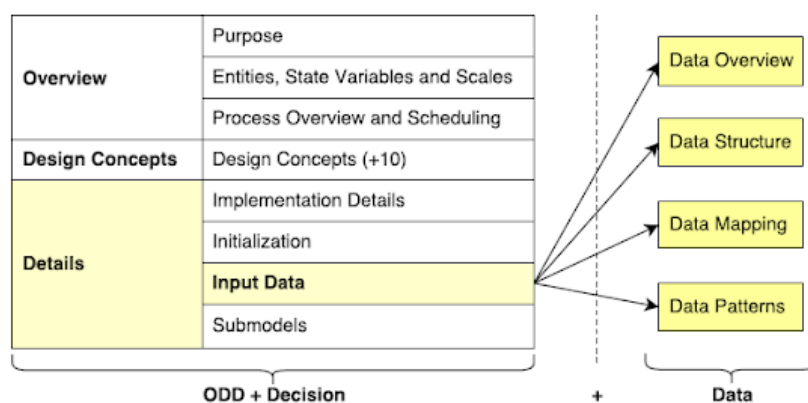


Figura 8 - Adição do campo Data ao protocolo ODD+D [4].

Renomeação do campo *Input Data* para *Data*, de maneira a apresentar melhor a relação com os dados como um todo, e não somente a entrada de dados, já que o uso de dados em modelos se tornaram mais diversos.

Adição do item de *Data Overview* ao elemento de *Data*, para que seja explicada a estrutura dos dados, se os mesmos estão disponíveis, qual sua origem e como eles foram coletados.

Adição do item de *Data Structure* ao elemento de *Data*, com o objetivo de apresentar quais são as variáveis, entidades e classes disponíveis nos dados, o que representam, em que formato estão, quais suas propriedades e como estão ligadas.

Adição do item de *Data Mapping* ao elemento de *Data*, para que seja esclarecido como os dados são ligados aos agentes e como cada agente faz uso dos dados disponibilizados para atingir o propósito do caso de estudo do modelo.

Adição do item de *Data Patterns* ao elemento de *Data*, para apresentar quais são os padrões existentes nos dados e como esses padrões afectam o comportamento dos agentes, além de como as alterações nos dados podem afectar os agentes.

Desta forma a versão mais atual do protocolo tem os seguintes elementos:

- *Overview;*
 - *Purpose;*
 - *Entities, state variables, and scales;*
 - *Process Overview and Scheduling;*
- *Design Concepts:*
 - *Theoretical and Empirical Background;*
 - *Individual Decision Making;*
 - *Learning;*
 - *Individual Sensing;*
 - *Individual Prediction;*
 - *Interaction;*
 - *Collectives;*
 - *Heterogeneity;*
 - *Stochasticity;*
 - *Observation;*
- *Details*
 - *Implementation Details;*
 - *Initialization;*
 - *Data;*
 - *Data Overview;*
 - *Data Structure;*
 - *Data Mapping;*
 - *Data Patterns;*
 - *Submodels.*

2.4. Trabalhos Relacionados

Um dos exemplos de trabalhos que se relacionam está em [29] no qual é apresentado o conceito de laboratório virtual para risco e impacto de dano. Neste trabalho é apresentado um laboratório virtual que funciona como um portal científico de fluxo de trabalho, onde é possível integrar diversos conhecimentos preexistentes.

Essa abordagem apresenta os conceitos de um repositório científico para compartilhamento de informações sobre pesquisas permitindo que esses conceitos sejam rastreados por meio do uso de proveniência dos dados, de maneira a identificar a origem dos elementos.

Uma iniciativa relevante é a plataforma CoMSES¹ que possui um repositório online para a consulta e partilha MBA. Nesta plataforma, numa lógica *open source* os modeladores e programadores, disponibilizam os modelos, com acesso às publicações, ao código das simulações e a descrição do MBA com os recursos as descrições ODD.

Como já mencionado, a relação entre MBA e ontologia é algo próximo, sendo alvo de estudo em diversos trabalhos. Entre estes trabalhos está [30], onde é apresentada a possibilidade de relacionar a abordagem de descrição ODD com uma visão ontológica. Entretanto o protocolo, quando usado, ainda se encontra na versão ODD. Ficam, portanto, alguns conceitos de fora da descrição dada pelo protocolo, tornando pouco profunda uma abordagem ontológica da representação das descrições.

Entretanto, com a evolução do protocolo até sua versão mais recente e face a importância de adicionar informações relativas ao código associado aos MBA e a proveniência dos modelos emergiu a necessidade de definição de uma ontologia mais abrangente. A existência de uma ontologia permite uma sistematização dos conceitos e a possibilidade de interpretação dos conceitos por humanos e máquinas. Com a existência de uma ontologia, surge, também, a possibilidade de criar ferramentas que permitem a rastreabilidade dos conceitos existentes nos modelos baseados em agentes descritos nas várias versões do protocolo ODD, garantindo assim o acesso às descrições de MBA escrita sem versões anteriores do protocolo.

2.5. Considerações Finais

Neste capítulo foram apresentados os conceitos de Web Semântica, com suas abordagens para a representação e o desenvolvimento da abordagem ontológica, bem como foram apresentados os conceitos de Modelos Baseados em Agentes, suas diferentes aplicações e foi contextualizado todo o histórico do protocolo de descrição. Por fim foram abordados os trabalhos que contém os conteúdos que se relacionam diretamente aos conceitos presentes no desenvolvimento deste trabalho.

¹ <https://www.comses.net>

Capítulo 3

Metodologia e Tecnologias

Neste capítulo, iremos abordar as tecnologias, técnicas e trabalhos gerais relacionados, com o trabalho apresentado neste documento.

3.1. Metodologia de desenvolvimento

Para a execução do projeto foi definida uma metodologia de desenvolvimento e elaborado um planeamento de acordo com as etapas, para garantir a boa execução do projeto e o seu sucesso.

3.1.1. Fluxograma de desenvolvimento

Para que existisse um melhor conhecimento da sequência lógica de execução das atividades, foi elaborado um fluxograma inicial que contém as atividades de desenvolvimento. Este fluxograma é apresentado na Figura 9.

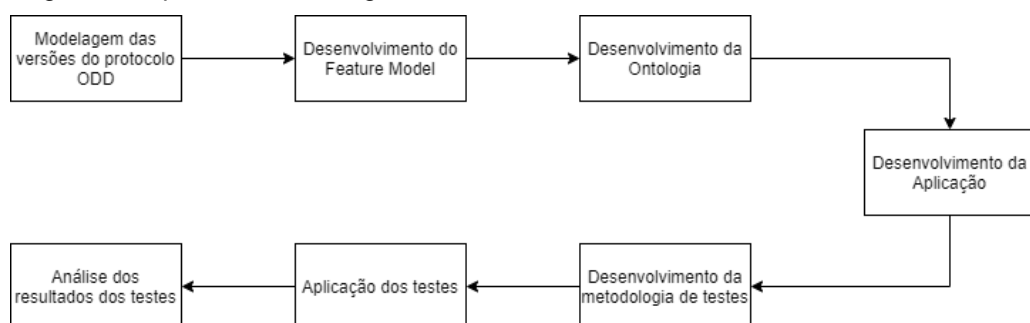


Figura 9 – Fluxograma de execução do projeto.

3.1.2. Cronograma

Com a definição das atividades a serem feitas, foi estabelecido um cronograma para o desenvolvimento do projeto, o qual pode ser visualizado na Tabela X.

Tabela 1 – Cronograma de execução do projeto.

Tarefa	Início	Fim
Definição das questões de investigação	01/10/2019	31/10/2019
Análise bibliográfica	01/11/2019	31/12/2019
Modelagem das versões do protocolo ODD	01/01/2020	31/01/2020
Desenvolvimento do Feature Model	01/02/2020	15/02/2020
Desenvolvimento da Ontologia	16/02/2020	31/03/2020
Desenvolvimento da Aplicação	01/04/2020	15/05/2020
Desenvolvimento da metodologia de testes	16/05/2020	15/09/2020
Aplicação e análise dos testes	16/09/2020	01/11/2020

3.2. Ontologia

Para o desenvolvimento da ontologia, foi feita uma análise do domínio com base numa revisão da literatura. Também foram utilizadas algumas ferramentas e métodos para representar os conceitos do domínio. Estas metodologias são descritas nesta secção.

3.2.1. Feature Model

O uso da abordagem de desenvolvimento de um feature model para o planeamento de uma ontologia é uma prática que em diversos casos faz sentido pois eles podem ser vistos como uma visualização da ontologia, devido a possuírem o mesmo nível de abstração, permitindo assim uma análise com maior facilidade dos elementos presentes na ontologia, permitindo uma compreensão melhor de como a ontologia é formulada e quais ajustes podem ser realizados, sendo assim uma escolha natural de abordagem para este projeto [31].

3.2.2. Protégé

O *Protégé* [32] é um software de código aberto que permite a criação e edição de ontologias, permitindo uma livre manipulação das ontologias e o suporte de uma grande comunidade académica, governamental e corporativa, sendo assim uma escolha lógica para o software de manipulação da ontologia.

3.2.3. Dublin Core

O Dublin Core² é um *schema* que permite a especificação de documentos, de maneira a permitir a integração à ontologia de uma especificação com a qualidade já avaliada por diversos profissionais especialistas, agregando assim maior qualidade a ontologia, e sendo uma escolha natural para o tratamento da especificação dos documentos na ontologia.

3.2.4. CodeMeta

O CodeMeta³ é um *schema* para a especificação de metadados sobre softwares, permitindo uma descrição completa das características presentes na aplicação, e da mesma forma do que com o Dublin Core, é uma integração que faz sentido a ontologia por trazer a qualidade de uma especificação já previamente avaliada por diversos especialistas.

3.3. Aplicação

Para o desenvolvimento da aplicação, diversas ferramentas foram necessárias, as quais estão apresentadas juntamente com a abordagem para a aplicação:

3.3.1. Ferramentas

Para a implementação da ferramenta de acesso a ontologia, era preciso realizar a escolha de alguma ferramenta capaz de realizar a transformação entre dados ontológicos e dados reconhecíveis para a apresentação, para isso foi escolhida a ferramenta RDFLib⁴, que é uma biblioteca em Python⁵, pois a mesma permite uma grande gama de integrações com diversos serviços de armazenamento, o que é interessante em casos de futuras trocas de armazenamento, além de possuir um serviço de armazenamento interno adaptável a uma grande gama de sistemas gerenciadores de banco de dados, o qual é ideal para uma fase de testes como a qual o projeto se encontra.

Por isso para o desenvolvimento da aplicação foi utilizada a abordagem de uma Interface de Programação de Aplicações (API) em arquitetura Transferência Representacional de Estado (REST), a qual foi desenvolvida com a linguagem de programação Python e utilizando o framework Flask⁶, esta aplicação permite o acesso a um servidor de arquivos, um gerenciador de usuários e um servidor dos dados ontológicos.

² <https://www.dublincore.org>

³ <https://codemeta.github.io>

⁴ <https://rdflib.readthedocs.io/en/stable>

⁵ <https://www.python.org>

⁶ <https://flask.palletsprojects.com/en/1.1.x>

Os dados ontológicos possuem o armazenamento na base de dados PostgreSQL⁷ por meio de um armazenamento adaptado de chave-valor, pois dentre todas as abordagens esta era a que melhor se adequava às necessidades do projeto, com as consultas aos dados sendo realizadas por meio do protocolo SPARQL, o qual é destinado a consultas em RDF.

Já os dados de utilizadores são armazenados em uma base de dados PostgreSQL com a estrutura de armazenamento tradicional. Por fim para a apresentação dos dados foi escolhida a abordagem de uma aplicação Angular, realizando a comunicação por meio de requisições HTTP.

Por fim a aplicação foi hospedada por meio da ferramenta Heroku⁸, disponibilizando o seu acesso por meio da Internet.

3.3.2. Estrutura

Para que a interação entre os diferentes elementos na aplicação pudesse ser melhor visualizada, estes elementos foram dispostos em um fluxograma, o qual é apresentado na Figura 10:

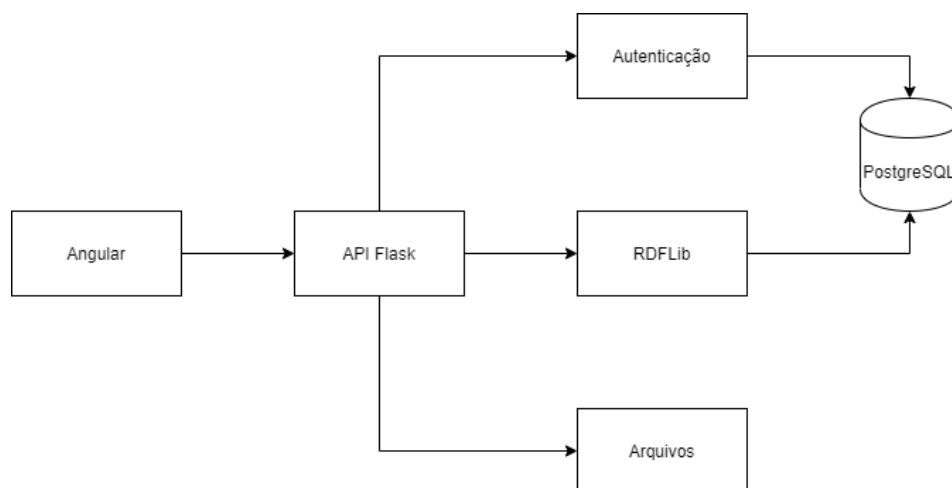


Figura 10 – Estrutura de comunicação de dados na aplicação.

3.4. Testes

Para a execução dos testes foram definidas, em primeiro lugar, as hipóteses a serem testadas. Na sequência desta definição, foi definida uma metodologia para a sua realização, conjuntamente com a escolha das ferramentas necessárias para isso.

⁷ <https://www.postgresql.org>

⁸ <https://www.heroku.com>

3.4.1. Metodologia de teste

Para desenvolver os testes e descrevê-los, a metodologia escolhida foi a apresentada em [33] mas de maneira adaptada para que se adequasse melhor às necessidades do projeto, por isso os elementos não serão citados diretamente, e nem todo o conteúdo será exposto diretamente, tendo em vista que alguns dos assuntos já foram expostos no texto.

3.4.2. Características dos testes

A seleção dos participantes para a realização dos testes foi dividida em 3 grupos, o primeiro grupo foi composto por alunos do primeiro ano do curso de licenciatura em Informática, pois esses usuários possuem um conhecimento sobre utilização de software, mas não possuem um conhecimento tão aprofundado sobre construção de software ou desenvolvimento de MBA, permitindo assim a avaliação do funcionamento básico da ferramenta para o aprendizado sobre MBA com os modelos presentes na ferramenta.

O segundo grupo já é composto por alunos do mestrado em engenharia de software, sendo com os usuários deste grupo possuindo um conhecimento maior de questões computacionais, os testes aplicados a este grupo permitem a avaliação da ferramenta para usuários que estejam em uma condição de fazer uma análise mais crítica sobre as possibilidades de aprendizado com a ferramenta.

Por fim o terceiro grupo é composto por peritos na área, os quais têm a possibilidade de avaliar precisamente quais as características da ferramenta atendem as necessidades destes profissionais, e de quais características carecem de alguma melhoria.

Devido às características para a execução dos testes a seleção dos usuários não poderia considerar características como a aleatoriedade na amostragem, ocorrendo por meio de uma amostragem quantitativa com todos os usuários que aceitaram participar do procedimento de testes.

Para a execução dos testes todo o procedimento precisou de um funcionamento a distância, pois devido a características de isolamento social outras alternativas não poderiam ser uma opção, obrigando que toda a coleção de informação ocorresse por meio da Internet.

Para a aplicação dos testes, foi formulada as hipóteses as quais fariam sentido verificar a experiência do usuário, para isso foram utilizados os objetivos do projeto, a partir disso foram definidas duas hipóteses a serem testadas:

- O usuário beneficia da utilização da ferramenta para melhor descrever um MBA?
- O usuário beneficia da utilização da ferramenta para melhor compreender um MBA?

Com essas duas hipóteses formuladas para os testes, foi definido qual seria a abordagem para executá-las. Para a primeira hipótese foi definida a abordagem de analisar o sucesso dos usuários durante o processo descritivo, além de pergunta-los se a abordagem lhes auxiliou e se possuem sugestões de melhoramento.

Já para a segunda hipótese, a abordagem foi de definir dois grupos, um que faria a análise por meio de um documento de texto, como habitualmente é realizada, e outro que faria a análise por meio da aplicação, após isso ambos os responderam um questionário sobre o MBA que leram, para assim analisar se uma das abordagens é mais proveitosa.

Além das questões voltadas para o teste das hipóteses, foram utilizadas questões que permitissem a percepção das características demográficas presentes nos grupos, com a coleta de informações como gênero, idade, e experiência profissional.

Para a coleta das informações de ambos os testes, foi utilizada a plataforma de formulários Google Forms⁹, que permite uma fácil disponibilização de formulários e coleta de dados. Para cada um dos testes foi utilizado um formulário, os quais podem ser visualizados nas listagens seguintes com os testes para inserção e visualização respectivamente.

Questionário para o teste de inserção:

- No modelo introduzido, os agentes tinham alguma capacidade de aprendizagem?
- Qual é o elemento do modelo usado para representar valores aleatórios?
- Para que elemento do modelo foi necessário usar uma imagem?
- Qual foi a maior dificuldade durante o teste?
- Teve algum problema no uso da aplicação?
- Que sugestões tem para a aplicação?

Questionário para os testes de interpretação:

- Qual o objetivo do modelo?
- Onde está disponível o modelo?
- Os agentes usam um o modelo próprio de decisão?
- Qual é a linguagem usada para a implementação do modelo?
- Qual o tipo de modelo?
- Quem inseriu os dados do modelo?

As análises foram feitas por meio de análises estatísticas com a ferramenta de planilha Google, devido a fácil integração com a ferramenta de formulários, essas análises juntamente com suas respectivas discussões se encontram no capítulo seguinte.

3.5. Considerações finais

Neste capítulo foram apresentadas todas as ferramentas e métodos utilizados durante a execução do processo, bem como a metodologia de desenvolvimento que foi utilizada, demonstrando como os elementos foram conduzidos para o desenvolvimento da ontologia e da aplicação e para a execução dos testes.

⁹ <https://docs.google.com/forms>

Capítulo 4

Resultados e Discussões

Neste capítulo serão apresentados os resultados obtidos a partir da execução deste projeto, conjuntamente com as discussões que sejam pertinentes a cada um dos resultados.

4.1. Ontologia

Como primeiro resultado está a ontologia, pois o desenvolvimento da ontologia foi a base para a execução do projeto.

4.1.1. *Objetivo*

O objetivo principal de nossa ontologia é apoiar os modeladores durante a atividade de elicitação num processo de desenvolvimento de MBA, bem como especificar os requisitos objetivos e os parâmetros de um MBA.

ODD+2D fornece uma base de conhecimento sobre os conceitos relacionados com o domínio. Esta ontologia foi criada para apoiar a elicitação e especificação dos MBA. A ontologia é um corpo de conhecimento para o domínio MBA na literatura. Deve harmonizar a terminologia de domínio as várias necessárias para a especificação de modelagem existentes para MBA e ajudar os modeladores a comunicar os seus modelos uns com os outros.

4.1.2. *Aquisição de conhecimento*

A aquisição de conhecimento é uma fase de desenvolvimento de ontologia, na qual podemos buscar documentos, livros, artigos, outras ontologias, taxonomias, normas internacionais e materiais sobre o domínio que possam ser utilizados como fonte de

conhecimento para a ontologia proposta. Como resultado desta fase, temos material suficiente para ser utilizado como fonte de conhecimento para a construção da ontologia.

Para compor a aquisição de conhecimento, consultamos os artigos de definição do protocolo ODD+2D, que consistem nos trabalhos [4], [26]–[28], que já foram apresentados anteriormente.

Esse conhecimento na sequência foi mapeado de maneira a existir uma coesão na representação que atenda a evolução do protocolo, além de facilitar a compreensão das representações necessárias para a sequência do trabalho, o diagrama pode ser visualizado nos Apêndices I, II e III.

Já para a obtenção de um modelo para ser utilizado como base no desenvolvimento da ontologia, foi desenvolvido o modelo de propriedades, que pode ser visualizado nas Figuras 11, 12 e 13, apresentando as características do ODD+2D utilizadas para o desenvolvimento da ontologia.

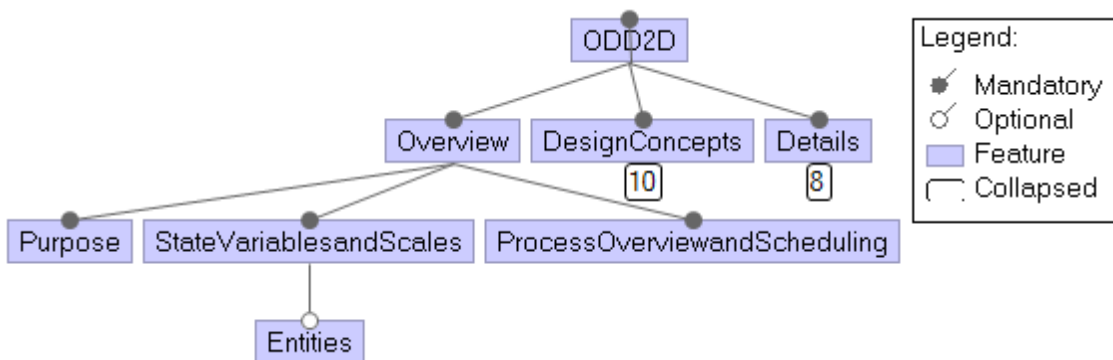


Figura 11 - Representação do elemento Overview na evolução de ODD ao ODD+2D no modelo de propriedades.

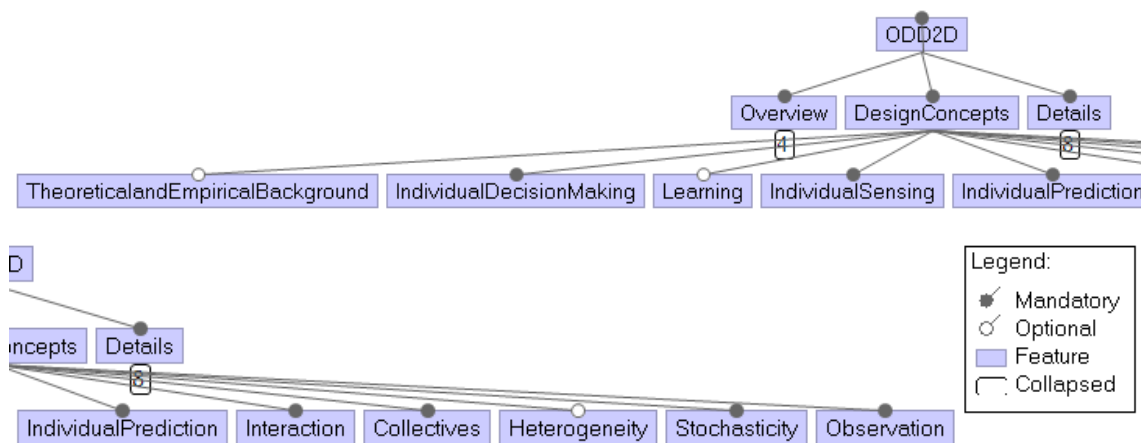


Figura 12 - Representação do elemento Design na evolução de ODD ao ODD+2D no modelo de propriedades.

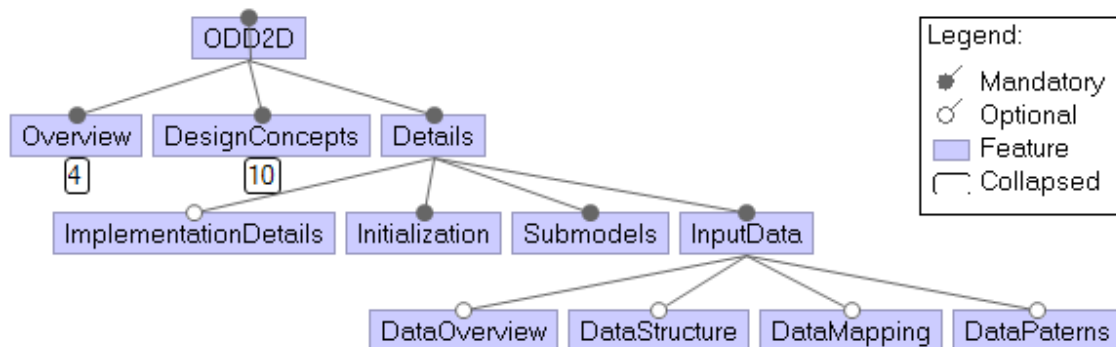


Figura 13 - Representação do elemento Details na evolução de ODD ao ODD+2D no modelo de propriedades.

Por fim, para acrescentar características relacionadas à documentação de arquivos de códigos, foram adicionados respectivamente elementos dos schemas Dublin Core e CodeMeta. Esses elementos foram selecionados de acordo com as necessidades apresentadas para a documentação de documentos ou códigos.

Os elementos selecionados no *Schema* Dublin Core foram: *Title*; *Creator*; *Contributor*; *Date*; *Subject*; *Language*; *Source*; *Rights*; *Publisher*; e *Identifier*. Esses elementos visam permitir que os utilizadores da ontologia possam vir a utilizá-la com a capacidade de descrever os documentos que fazem parte do MBA descrito.

Por sua vez os elementos adicionados a partir do CodeMeta foram os seguintes: *URL*; *Maintainer*; *Version*; *Reference Publication*; *Programming Language*; *File Format*; *File Size*; *License*; e *Read Me*. Esses elementos foram selecionados por permitirem uma descrição das abordagens utilizadas na codificação do MBA, de uma forma que seja abrangente ao mesmo tempo que permite ao utilizador detalhar o seu trabalho.

4.1.3. Conceptualization

A conceptualização é uma fase na qual capturamos os conceitos, relações e propriedades da ontologia com base nas questões de competência. Na fase de conceptualização, descrevemos os conceitos e relações da ontologia proposta. Todo conceito identificado na aquisição de conhecimento faz parte do ODD+2D. Primeiramente, apresentamos os conceitos estruturados em um glossário com suas respectivas descrições. A seguir, definimos a relação entre os conceitos. Por fim, apresentamos as propriedades do ODD+2D. Usamos Protege com o plugin DataViz para criar um modelo gráfico.

4.1.3.1. Conceitos

Nesta seção, vamos definir os termos usados antes de desenvolver uma hierarquia de classes e as propriedades dos conceitos.

- *Agent*: é um indivíduo concreto existencialmente independente (entidades reais

possuidoras de uma identidade única) que não possui partes temporais e persiste no tempo mantendo sua identidade [34].

- *Title*: conceito do *schema* Dublin Core para representar o título do documento
- *Creator*: conceito do *schema* Dublin Core para representar os autores do documento
- *Contributor*: conceito do *schema* Dublin Core para representar os contribuidores do documento
- *Date*: conceito do *schema* Dublin Core para representar a data do documento
- *Subject* - conceito do *schema* Dublin Core para representar o tema do documento
- *Language*: conceito do *schema* Dublin Core para representar a língua do documento
- *Source*: conceito do *schema* Dublin Core para representar a fonte de disponibilização do documento
- *Rights*: conceito do *schema* Dublin Core para representar os direitos relacionados ao documento
- *Publisher*: conceito do *schema* Dublin Core para representar a editora do documento
- *Identifier*: conceito do *schema* Dublin Core para representar o identificador do documento
- *URL*: conceito do *schema* CodeMeta para representar a URL do código
- *Maintainer*: conceito do *schema* CodeMeta para representar o mantenedor do código
- *Version*: conceito do *schema* CodeMeta para representar a versão do código
- *Reference Publication*: conceito do *schema* CodeMeta para representar a publicação de referência do código
- *Programming Language*: conceito do *schema* CodeMeta para representar a linguagem de programação do código
- *File Format*: conceito do *schema* CodeMeta para representar o formato do arquivo do código
- *File Size*: conceito do *schema* CodeMeta para representar o tamanho do arquivo do código
- *License*: conceito do *schema* CodeMeta para representar a licença do código
- *Read Me*: conceito do *schema* CodeMeta para representar o arquivo de instruções do código
- *Overview*: conceito do protocolo ODD+2D para representar a introdução do MBA
- *Purpose*: conceito do protocolo ODD+2D para representar a proposta do MBA
- *State Variables and Scales*: conceito do protocolo ODD+2D para representar as variáveis e escalas do MBA
- *Entities*: conceito do protocolo ODD+2D para representar as entidades do MBA

- *Process Overview and Scheduling*: conceito do protocolo ODD+2D para representar a introdução ao funcionamento das entidades do MBA
- *Design Concepts*: conceito do protocolo ODD+2D para representar as escolhas de desenho do MBA
- *Theoretical and Empirical Background*: conceito do protocolo ODD+2D para representar os conceitos gerais, teóricos e as hipóteses do MBA
- *Individual Decision Making*: conceito do protocolo ODD+2D para representar as escolhas de decisões dos agentes do MBA
- *Learning*: conceito do protocolo ODD+2D para representar o processo de aprendizado dos agentes do MBA
- *Individual Sensing*: conceito do protocolo ODD+2D para representar as habilidades de sensibilidade dos agentes do MBA
- *Individual Prediction*: conceito do protocolo ODD+2D para representar as habilidades de predição dos agentes do MBA
- *Interaction*: conceito do protocolo ODD+2D para representar as capacidade de interação dos agentes do MBA
- *Collectives*: conceito do protocolo ODD+2D para representar se os agentes são capazes de agir como grupos do MBA
- *Heterogeneity*: conceito do protocolo ODD+2D para representar a heterogeneidade dos agentes do MBA
- *Stochasticity*: conceito do protocolo ODD+2D para representar a aleatoriedade do MBA
- *Observation*: conceito do protocolo ODD+2D para representar a o que é coletado de dados dos agentes do MBA
- *Details*: conceito do protocolo ODD+2D para representar os detalhes no desenvolvimento do MBA
- *Implementation Details*: conceito do protocolo ODD+2D para representar os detalhes presentes na implementação do MBA
- *Initialization*: conceito do protocolo ODD+2D para representar a inicialização do MBA
- *Submodels*: conceito do protocolo ODD+2D para representar os submodelos do MBA
- *Data*: conceito do protocolo ODD+2D para representar os dados do MBA
- *Data Overview*: conceito do protocolo ODD+2D para representar a introdução dos dados do MBA
- *Data Structure*: conceito do protocolo ODD+2D para representar a estrutura dos dados do MBA
- *Data Mapping*: conceito do protocolo ODD+2D para representar o mapeamento dos dados do MBA
- *Data Patterns*: conceito do protocolo ODD+2D para representar o padrão dos

4.1.3.2. Relacionamentos

Uma vez definidas as classes e propriedades destas classes, vamos criar a hierarquia entre as classes da ontologia. Os relacionamentos definidos se dão na modelagem do protocolo ODD+2D, pois o mesmo exige a representação relacional dos conceitos que são atributos de outros conceitos. A compreensão dessa representação fica mais clara por meio da Figura seguinte, a qual permite visualizar de maneira mais profunda os relacionamentos presentes nos conceitos, e da próxima seção que aborda os atributos. A Figura 14 representa o relacionamento entre o primeiro nível dos conceitos o protocolo ODD, ODD+D, ODD+2D, Dublin Core e Code Meta. As Figuras 15, 16, 17 e 18 mostram a hierarquia entre as subclasses.

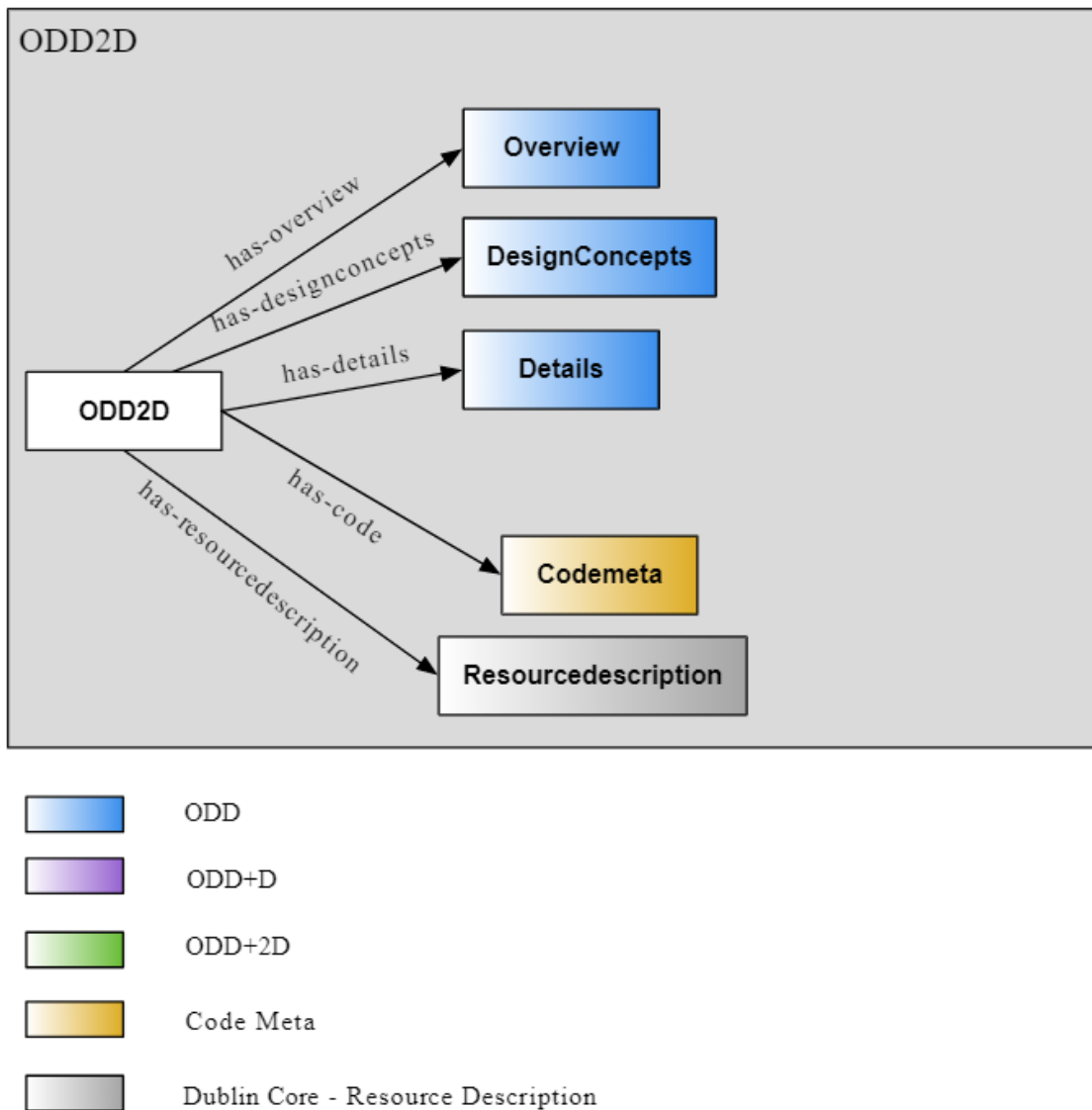
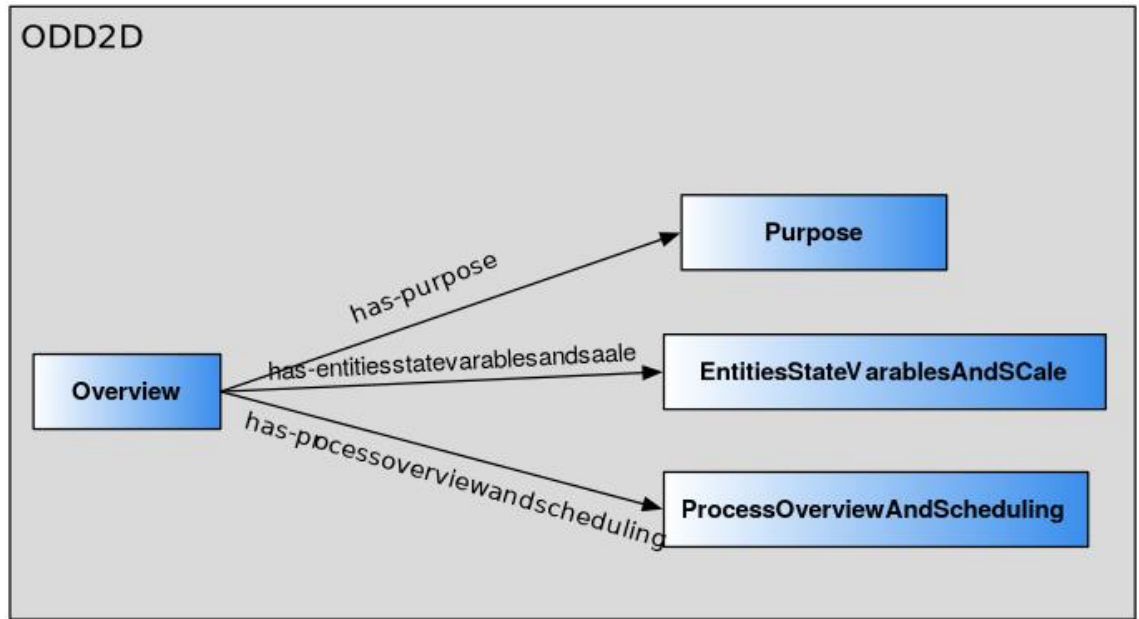


Figura 14 – Representação em primeiro nível da ontologia.



 ODD

Figura 15 – Representação do elemento *Overview* na ontologia.

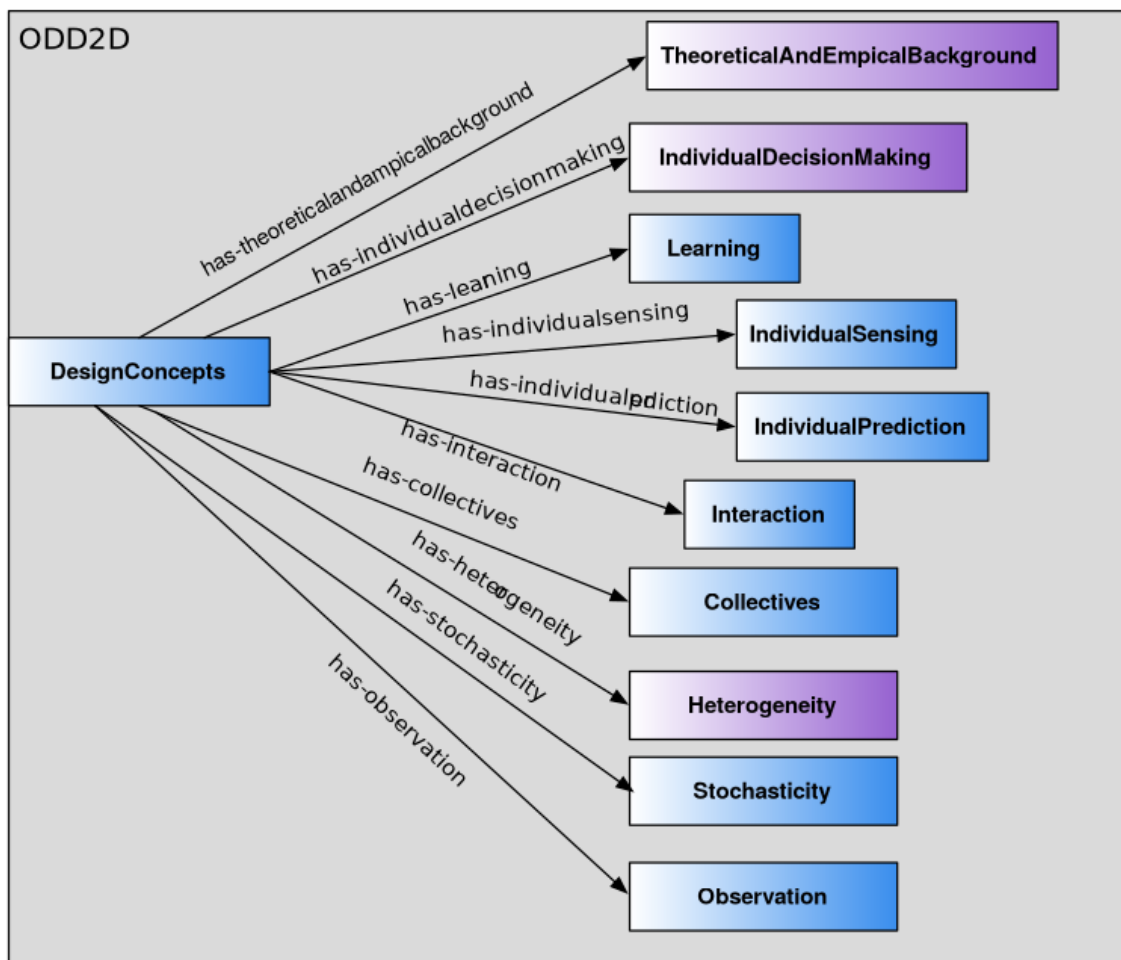


Figura 16 – Representação do element *Design Concepts* na ontologia.

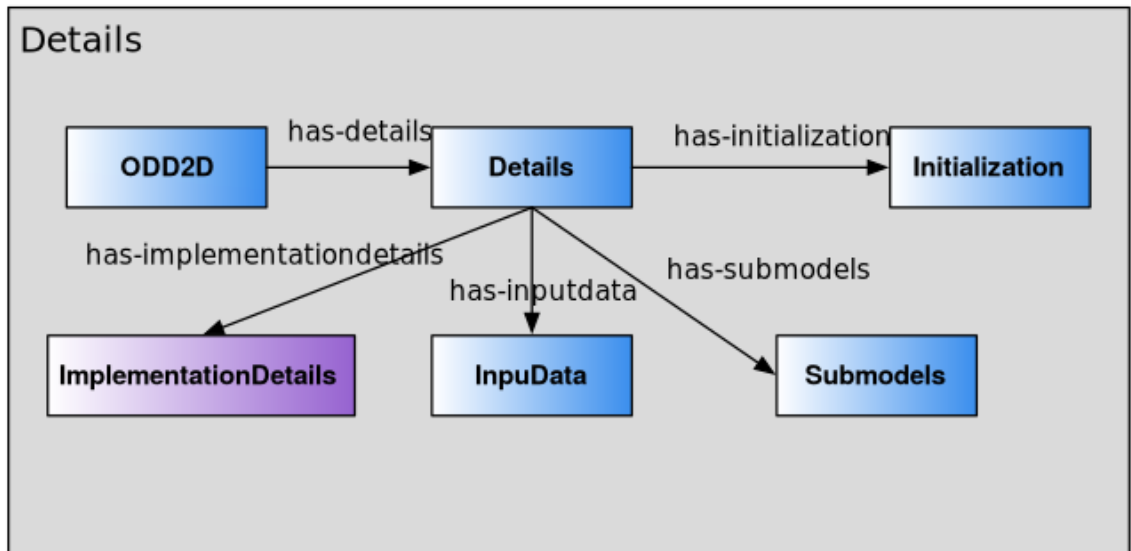


Figura 17 – Representação do elemento *Details* na ontologia.

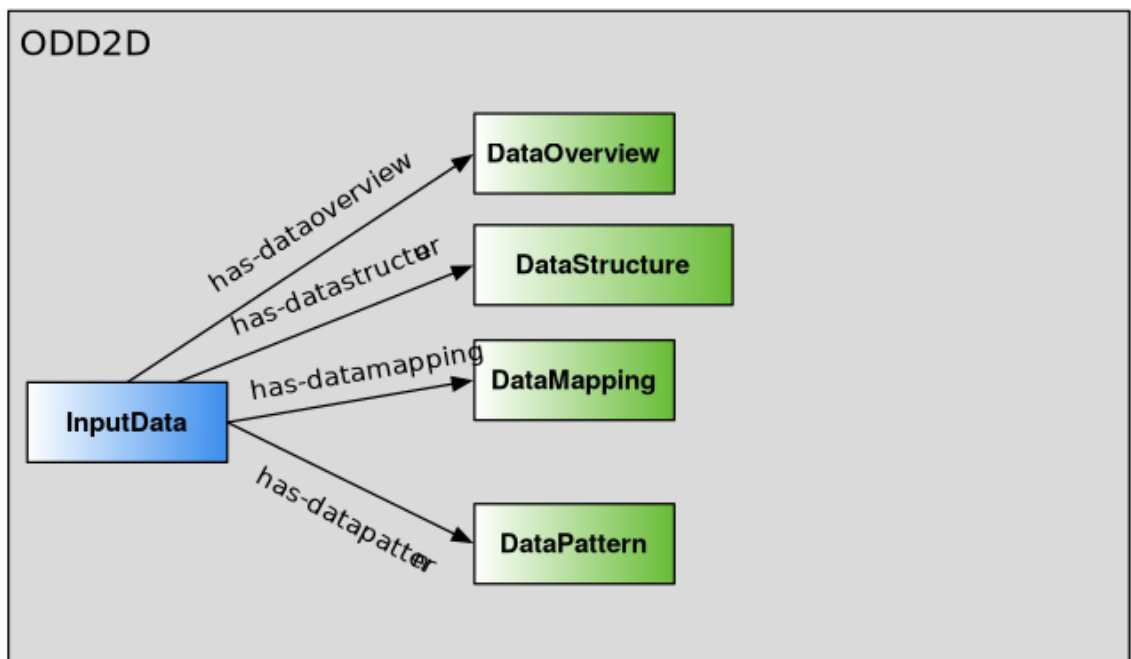


Figura 18 – Representação do elemento *Data* na ontologia.

4.1.3.3. Atributos

As ontologias também apresentam atributos. Eles são as propriedades dos conceitos das ontologias, bem como seus tipos de valores, que podem conter o Identificador de Recursos Internacionalizado (IRI). A Tabela 2 apresenta os atributos de ODD+2D.

Tabela 2 – Os conceitos e seus atributos na ontologia.

Conceito	Atributo	Tipo de Valor
ODD+2D	<i>Overview</i>	Literal ou IRI
Overview	<i>Purpose</i>	Literal ou IRI
<i>Overview</i>	<i>State Variables and Scales</i>	Literal ou IRI
<i>State Variables and Scales</i>	<i>Entities</i>	Literal ou IRI
<i>Overview</i>	<i>Process Overview and Scheduling</i>	Literal ou IRI
ODD+2D	<i>Design Concepts</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Theoretical and Empirical Background</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Individual Decision Making</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Learning</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Individual Sensing</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Individual Prediction</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Interaction</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Collectives</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Heterogeneity</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Stochasticity</i>	Literal ou IRI
<i>Design Concepts</i>	<i>Observation</i>	Literal ou IRI
ODD+2D	<i>Details</i>	Literal ou IRI
<i>Details</i>	<i>Implementation Details</i>	Literal ou IRI
<i>Details</i>	<i>Initialization</i>	Literal ou IRI
<i>Details</i>	<i>Submodels</i>	Literal ou IRI
<i>Details</i>	<i>Data</i>	Literal ou IRI
<i>Data</i>	<i>Data Overview</i>	Literal ou IRI
<i>Data</i>	<i>Data Structure</i>	Literal ou IRI
<i>Data</i>	<i>Data Mapping</i>	Literal ou IRI
<i>Data</i>	<i>Data Patterns</i>	Literal ou IRI

4.1.4. Implementação

Dado que nosso objetivo era desenvolver uma ontologia a fim de auxiliar na elicitaco e especificaco de MBA e torn-la (re)utilizvel pela comunidade de modeladores, avaliamos a ontologia por usabilidade, utilidade e facilidade de uso. Para completar, podemos ilustrar o uso da ontologia e compar-la com trabalhos relacionados. Para a verificaco, podemos criar uma tabela com as questes de competncia e conceitos de ontologia, relates e axiomas como

respostas. A validação pode ser realizada por especialistas de domínio e instanciação de um sistema do mundo real com os elementos da ontologia. Para os critérios de usabilidade, utilidade, facilidade de uso, podemos usar um estudo de caso e uma pesquisa para atingir esses critérios. O resultado desta fase é a documentação das avaliações realizadas.

4.1.5. Aplicação

A ferramenta resultante deste trabalho visa a inserção, pesquisa e visualização de MBA, e para atender as diferentes necessidades se divide em duas, sendo uma ferramenta com o foco na interação com máquinas (em forma de uma API pública para acesso e compartilhamento de dados) e outra com humanos (um cliente Web que pode ser acessado pelos utilizadores com maior facilidade).

A API desenvolvida se caracteriza por disponibilizar 4 recursos: Usuário (acessível por meio do */user* e */user/<id>*), Upload (acessível por meio do */file*), Download (acessível por meio do */<type>/<file>*) e Ontologia (acessível por meio do */rdf*, */rdf/<id>* e */rdf/query*), esses elementos visam atender as necessidades desde a criação e autenticação de usuários, o upload e conversão de arquivos em um acesso de Localizador Uniforme de Recursos (URL), a disponibilização do acesso dos arquivos identificados de por meio da URL, e o acesso ontológico aos dados, permitindo que as consultas ontológicas possam ser feitas livremente.

O recurso Usuário permite o método *post* para a execução dos métodos *signin* e *login*, para respectivamente o cadastro e acesso dos utilizadores. Já o recurso *upload* permite o método *post file*, para a execução do envio de um arquivo e disponibilização de uma URL de acesso ao arquivo, enquanto o recurso *Download* permite o acesso ao arquivo enviado por meio da URL de acesso retornada no método de envio.

Por sua vez o recurso Ontologia apresenta a parte lógica da aplicação, permitindo por meio dos métodos *post*, *put*, *get* e *delete* a criação, alteração, consulta e exclusão de dados, entretanto a parte de maior destaque permitida pela aplicação é a execução de consultas ontológicas aos dados, permitindo de acordo com os conceitos da ontologia anteriormente apresentados estabelecer relação entre os diferentes modelos, permitindo assim a rastreabilidade dos dados de diferentes MBA.

Já a aplicação Web permite o acesso dos clientes aos mesmos conceitos presentes na API, mas lhes permitindo um acesso facilitado aos conceitos, em decorrência da vantagem de apresentar os dados de maneira gráfica. Uma das vantagens que o utilizador tem na aplicação Web é o uso de formulários para criação da descrição, pois devido a grande quantidade de informações necessárias sobre o MBA, o uso de um formulário estruturado de acordo com as etapas de descrição, e com instruções detalhadas dos conceitos necessários em cada um dos campos, permite uma facilidade maior no desenvolvimento da descrição.

Mas a maior vantagem da aplicação é o seu uso para visualização e rastreabilidade de modelos, pois a aplicação fornece uma possibilidade de execução das consultas de rastreabilidades por meio de um formulário, tornando muito mais intuitivo a execução das

consultas. Além disso, a aplicação permite a apresentação de um grafo que detalha quais são os conceitos apresentados no modelo, juntamente com o detalhamento seccionado dos conceitos do MBA, deixando mais claro ao leitor o que está presente em cada etapa.

Por isso foi desenvolvido um modelo que permite tanto a facilidade do acesso de humanos por meio de uma aplicação Web, quanto o acesso a máquinas por meio de uma API para a integração da ontologia em outras aplicações do meio de MBAs, permitindo assim que a ontologia seja acessível a todos que tiverem interesse, e assim popularizar o modelo de descrição de MBAs, facilitando a divulgação dos conhecimentos desenvolvidos.

4.2. Processo

Nesta seção, explicamos como usar a ontologia proposta nesta tese, tanto em sua forma de aplicação como de API. Portanto, propomos o processo ODDEspec que instancia o processo de especificação de uma instância de um MBA. A instância criada define os artefatos de entrada e saída de cada atividade e como proceder nas atividades, utilizando a ontologia ODD+2D. Este processo visa auxiliar o modelador no processo de especificação e documentação de um MBA. Ele permite que um modelador elicite e especifique um MBA.

4.2.1. Visão geral do processo ODDEspec

Como primeira etapa do processo, está a inserção dos dados de um modelo, essa execução é apresentada na Figura 19, e demonstra os principais processos para que os dados de um MBA sejam inseridos.

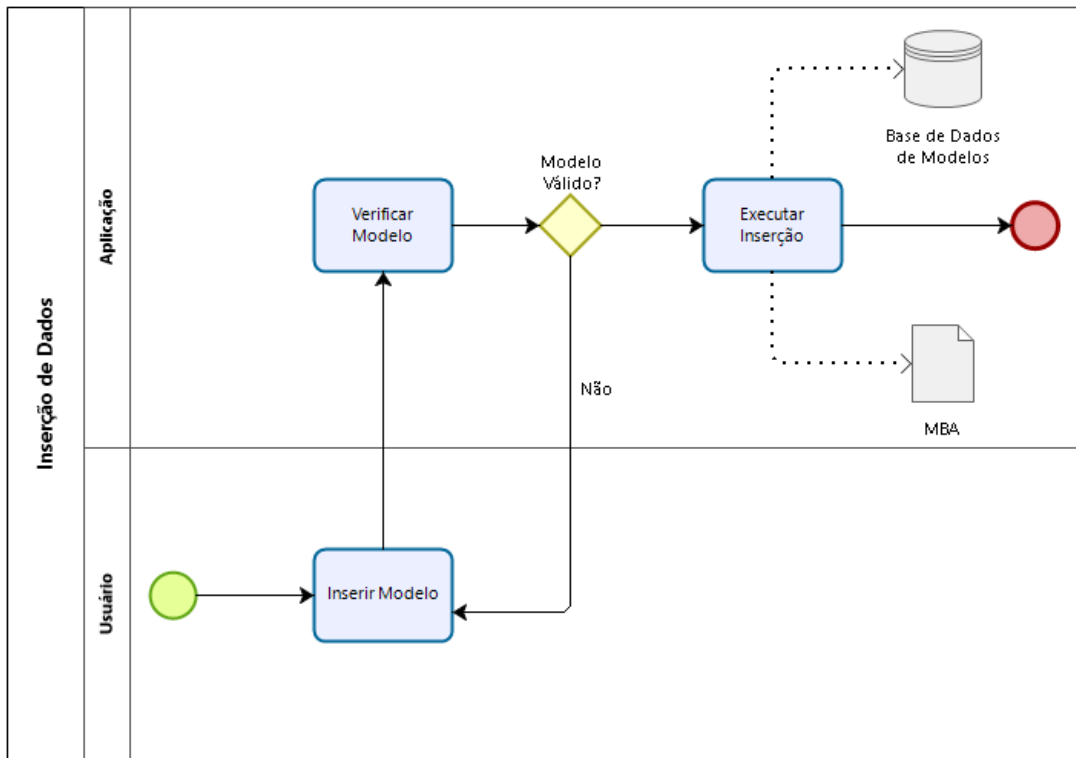
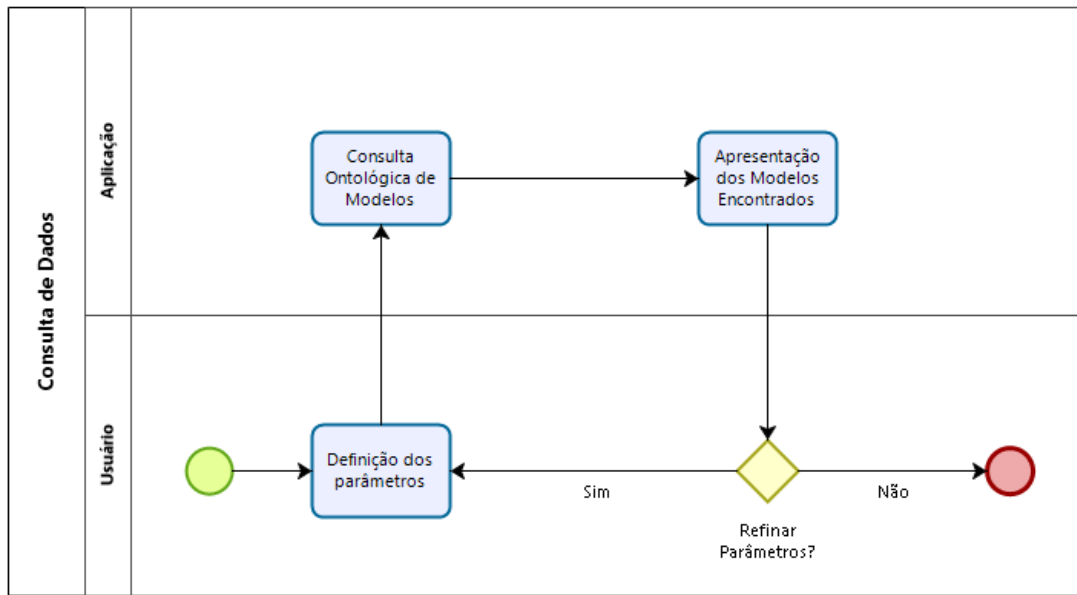


Figura 19 – A inserção de dados de um MBA.

O processo se inicia com o preenchimento do formulário por parte do utilizador, o qual conta com perguntas de apoio para a tomada de decisão do conteúdo a ser preenchido, essas perguntas têm como base as perguntas de apoio para preenchimento do protocolo ODD+2D nos trabalhos [4], [28]. Na sequência a aplicação verifica se o utilizador preencheu todos os campos obrigatórios, retornando para o preenchimento do utilizador caso algum campo necessite ser preenchido, caso contrário segue para a inserção dos dados na ontologia.

O segundo processo para utilizadores na aplicação é a consulta de dados, o qual é apresentado na Figura 20.

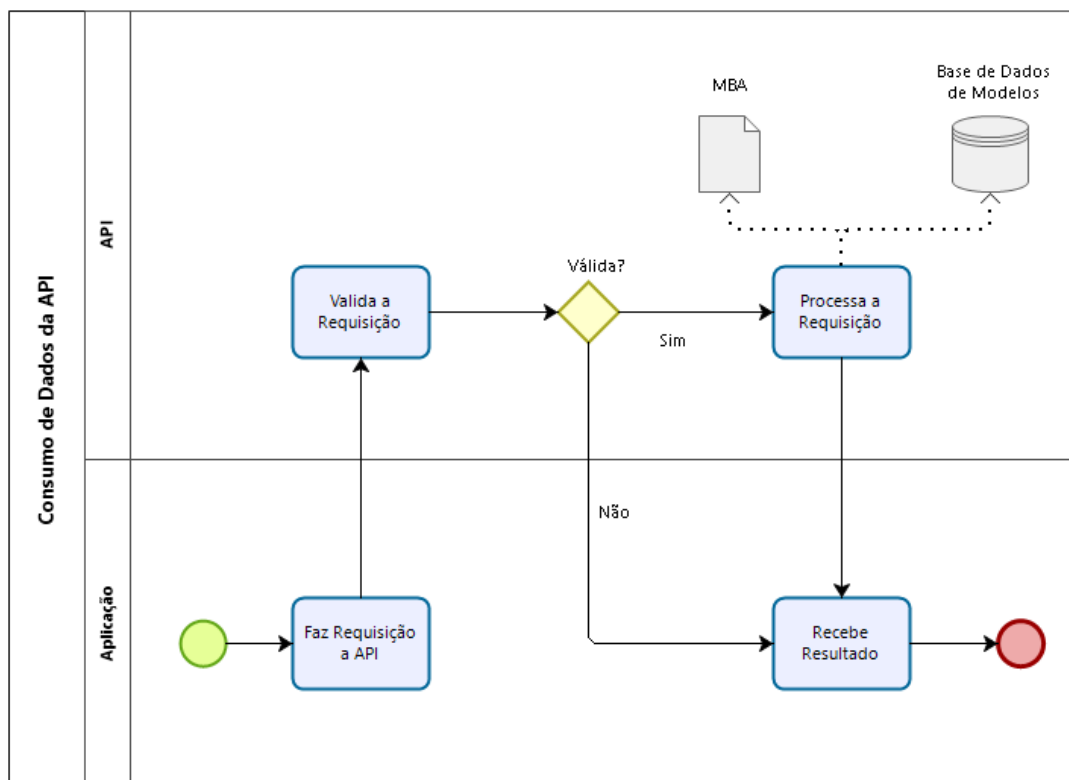


Powered by
bizagi
Modeler

Figura 20 – Consulta de dados de um MBA.

A consulta é iniciada pela definição de parâmetros, os quais podem não existir para uma consulta geral, podem ser definidos de acordo com o identificador de um modelo para uma consulta unitária, ou podem ser definidos de acordo com os atributos de modelos, para a rastreabilidade de modelos. De acordo com os parâmetros definidos a consulta ontológica é realizada pela aplicação, que na sequência exibe os dados, sendo a opção de consulta por um modelo específico a que possui mais opções de exibição, por apresentar um detalhamento de todos os atributos e um diagrama que apresenta quais itens do protocolo foram utilizados na especificação do modelo.

Por fim, ainda existe a opção de interação com a ontologia por meio de um acesso a API, que visa facilitar o acesso de máquinas a ontologia. O processo de acesso a API é o mais básico, pois todo acesso a aplicação é uma abstração do acesso à API. a lógica de processo pode ser vista no diagrama da Figura 21:



Powered by
bizagi
Modeler

Figura 21 – Acesso a API como intermédio para os MBAs.

A aplicação que deseja consumir dados da API deve enviar uma requisição de acordo com o conteúdo que deseja, essa requisição então é validada, caso não seja válida essa informação é retornada para a aplicação, e caso seja os dados são então processados de acordo com as características da requisição, podendo fazer um processamento ontológico, de acesso a arquivos ou de acesso de usuário.

4.3. Avaliação

O trabalho apresentado em [35] possui uma classificação das abordagens de avaliação de ontologias. Ele defende que a maioria das abordagens de avaliação se enquadram em uma das seguintes categorias:

- Comparar a ontologia a um “padrão ouro” (que pode ser uma ontologia);
- Utilizar a ontologia em uma aplicação e avaliar os resultados;
- Comparações com uma fonte de dados (por exemplo, uma coleção de documentos) sobre o domínio a ser coberto pela ontologia;
- Humanos que tentam avaliar o quão bem a ontologia atende a um conjunto de critérios, padrões, requisitos predefinidos, etc.

Com base nisso, pretendemos avaliar nossa ontologia usando critérios: usabilidade, utilidade e facilidade de uso.

4.3.1. Uma ferramenta de inserção, pesquisa e visualização de MBA

Para realizar os testes na aplicação, foi desenvolvido um protocolo de teste como apresentado na seção 4 do capítulo 3, de maneira a avaliar a aplicação por meio de estudantes que se enquadrem no perfil de pessoas que poderão fazer uso da aplicação no futuro, avaliando o desempenho da aplicação para inserção pesquisa e visualização de MBA, de maneira a torná-la melhor aos utilizadores.

4.3.2. Usabilidade, Utilidade e facilidade de uso

Para avaliar a aplicação, foi aplicada a abordagem de testes apresentada no capítulo anterior, avaliando os usos para inserção e interpretação dos dados com estudantes de na área de computação de graduação e mestrado, de maneira a abranger diferentes perfis de utilizadores, para agregar uma diversidade de opiniões sobre os elementos a serem analisados.

4.3.3. Resultados e discussões

Para a avaliação da inserção foram feitas 6 perguntas, sendo 2 sobre o MBA inserido e 4 sobre o processo de inserção. Nas 2 perguntas relacionadas ao MBA inserido, os utilizadores que realizaram os testes apresentaram como tiveram dificuldades em interpretar o modelo para inserção, pois na pergunta “No modelo introduzido, os agentes tinha alguma capacidade de aprendizagem?”, apenas 40% dos estudantes responderam corretamente, enquanto na pergunta “Qual é o elemento do modelo usado para representar valores aleatórios? ”, apenas 20% responderam corretamente.

Apesar de o modelo para inserção escolhido para ser utilizado pelos alunos, ainda se tratava de um documento com 24 páginas e um conteúdo extremamente técnico, dificultando a fácil compreensão necessária para um teste que se propunha ter entre 30 minutos e 1 hora. Entretanto esses elementos já apresentam indicadores de como a descrição de MBA requisita uma especialidade muito grande, não sendo algo de fácil adaptação para alunos de graduação, mesmo esses já tendo grande conhecimento em computação.

Por sua vez em relação as 4 perguntas relacionadas com a inserção, os utilizadores apresentaram visões mais distintas na forma de uso da aplicação, na pergunta “Para que elemento do modelo foi necessário usar uma imagem? ”, com as opções “Process Overview and Scheduling”; “State Variables and Scales”; e “Purpose”, 20% dos estudantes optaram por inserir em todos os campos, outros 20% em 2 campos, 40% em 1 campo, e 20% optaram por

não inserir imagens.

Já para a pergunta “Qual foi a maior dificuldade durante o teste?” 80% dos estudantes responderam que a interpretação foi seu maior problema, com 20% dos utilizadores tendo maiores dificuldades com a aplicação, e na pergunta “Teve algum problema no uso da aplicação?”, todos os estudantes apresentaram dificuldades em entender o que deveria ser preenchido em um campo, 80% não souberam escolher o tipo de dado e 60% não souberam fazer o upload do arquivo corretamente. Por fim para a pergunta de sugestão de melhoria os utilizadores indicaram a necessidade de uma maior segmentação da aplicação, para que seja mais clara em cada campo a ser preenchido, e pela indicação individual de campos obrigatórios não preenchidos.

Esses resultados indicam que em decorrência de os usuários terem tido dificuldades na interpretação do modelo, acabaram por utilizar abordagens diferentes para o preenchimento do formulário. Isso soma-se ao fato de necessitarem interpretar um texto corrido para segmentá-lo de acordo com as informações necessárias, faz com que a interpretação pessoal crie resultados diferentes no preenchimento do formulário.

Além disso o fato de a ontologia requerer muitas informações, com o formulário apresentando 44 campos para preenchimento, isso fez com que os usuários apresentassem dificuldades em entender o que deveria ser preenchido no campo, e como deveria ser preenchido, chegando ao ponto de apresentar dúvidas se os usuários fariam o upload do conteúdo em arquivo como texto, URL, arquivo de texto, imagem ou tabela. Além disso a grande quantidade de informação a ser preenchida acabou por causar dúvidas nos utilizadores em relação ao que ainda não havia sido preenchido, com os utilizadores sugerindo uma segmentação maior dos conteúdos apresentados na tela.

Para o teste de interpretação também foram feitas 6 perguntas, entretanto nesse caso todas as perguntas foram sobre o MBA, para verificar se os estudantes conseguiriam extrair as informações sobre o MBA lhes apresentado na aplicação com rapidez. Para as perguntas “Qual o objetivo do modelo?”, “Onde está disponível o modelo?”, “Quem inseriu os dados do modelo?”, e “Os agentes usam um modelo próprio de decisão?”, todos os utilizadores obtiveram a resposta correta, já para as perguntas “Qual é a linguagem usada para a implementação do modelo?” e “Qual o tipo de modelo?” 20% dos estudantes apresentaram uma resposta incorreta.

Para melhor compreender o sucesso dos utilizadores nos testes com a aplicação, foi desenvolvido um gráfico que apresenta a taxa de sucesso tanto no teste de inserção quanto no teste de interpretação, sendo o primeiro com uma taxa de sucesso de 100% (pois todos os utilizadores conseguiram inserir os dados corretamente) e o segundo 93% (pois este foi o desempenho dos utilizadores nas questões), como apresentado na Figura 22.

A Taxa de Sucesso nos Testes de Inserção e Interpretação

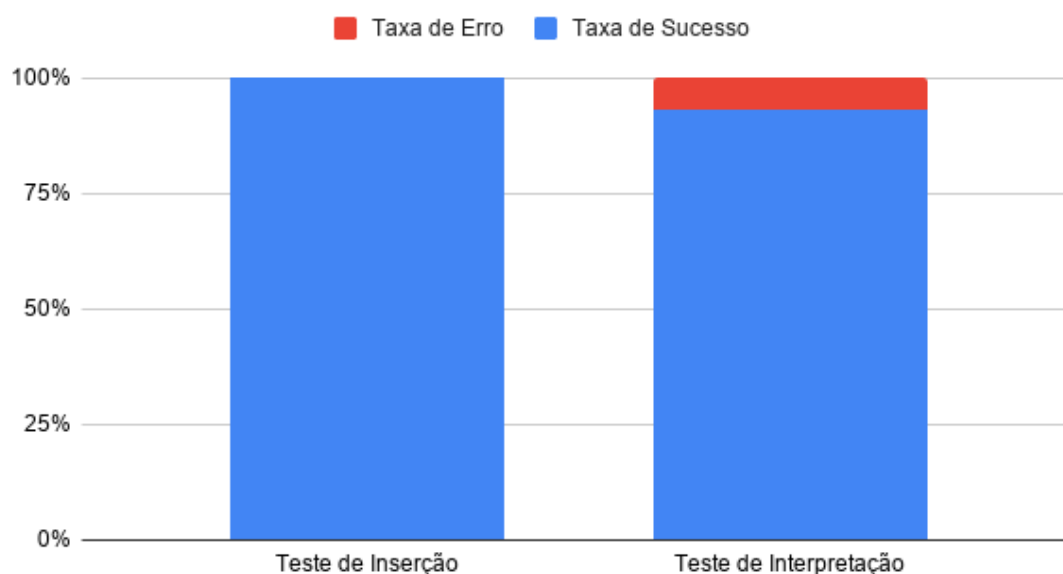


Figura 22 - Taxa de sucesso nos testes de inserção e interpretação.

Com esses resultados podemos perceber como a aplicação permite aos utilizadores um acesso mais facilitado aos conceitos do MBA, principalmente quando comparados aos estudantes que realizaram a interpretação por meio do documento em texto no teste de inserção. A aplicação lhes permitiu serem objetivos nos conceitos do MBA, devido a apresentação segmentada de cada um dos conceitos em representações específicas da ontologia.

4.4. Considerações Finais

Neste capítulo foram apresentados os resultados obtidos por meio deste trabalho, desde os resultados do desenvolvimento da ontologia, passando pelos resultados do desenvolvimento da aplicação, e a avaliação de usuários sobre a aplicação. No capítulo seguinte serão abordadas as conclusões resultantes do trabalho.

Capítulo 5

Conclusão e Trabalhos Futuros

5.1. Conclusão

Neste documento, apresentamos questões relacionadas com a especificação dos Modelos Baseados em Agentes. Além de discutir os problemas relacionados com esta especificação, foi apresentada uma proposta de representação, baseada numa ontologia, do protocolo ODD+2D, usado para descrever MBA, e das suas extensões ODD+D e ODD+2D. Esta solução permite tratar de maneira computacional o conteúdo até então somente descritivo.

A partir desta ontologia, torna-se possível realizar a rastreabilidade dos MBA de acordo com todas as suas características, considerando a proveniência dos modelos e dos seus dados. A representação dos dados do modelo como *linked data* possibilita a integração com a Web Semântica. o desenvolvimento dos MBA.

A proposta foi aplicada com sucesso num protótipo desenvolvido para criar e consultar modelos baseados em agentes usando a ontologia. Este protótipo integra uma aplicação Web e uma API, por forma a permitir o maior acesso possível.

Este protótipo foi submetido a testes por usuários de diferentes perfis. Estes testes permitiram identificar potencialidades e limitações da aplicação. Foram determinados pontos a serem melhorados, garantindo assim que a aplicação tenha um melhor uso pelos seus utilizadores finais.

A partir disso este trabalho disponibiliza um *framework* o qual permite tanto a utilizadores humanos quanto por meio da integração de máquinas a documentação de Modelos Baseados em Agentes, permitindo assim um acesso de uma comunidade de desenvolvimento a este trabalho, e permitindo trabalhos futuros.

5.2. Trabalhos Futuros

Para continuar a desenvolver a aplicação criada, a próxima etapa deste trabalho consiste numa fase de utilização do protótipo por utilizadores peritos do domínio para criar uma base consistente de modelos e recolher dados de utilização. Com a conclusão deste trabalho, foi iniciado o processo de desenvolvimento de um artigo sobre o mesmo, o qual está em processo de finalização para ser submetido para a conferência COMPSAC 2021: Intelligent and Resilient Computing for a Collaborative World ¹⁰, a qual permitirá acrescentar ao trabalho uma perspectiva diferente de avaliação.

¹⁰ <https://ieeecompsac.computer.org/2021/>

Bibliografia

- [1] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. SUPPL. 3, pp. 7280–7287, 2002.
- [2] M. C. González, C. A. Hidalgo, and A. L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [3] N. Djurdjevac Conrad, L. Helfmann, J. Zonker, S. Winkelmann, and C. Schütte, "Human mobility and innovation spreading in ancient times: a stochastic agent-based simulation approach," *EPJ Data Sci.*, vol. 7, no. 1, p. 24, 2018.
- [4] A. Laatabi, N. Marilleau, T. Nguyen-Huu, H. Hbid, and M. A. Babram, "ODD+2D: An ODD based protocol for mapping data to empirical ABMs," *JASSS*, vol. 21, no. 2, 2018.
- [5] N. Gibbins, S. Harris, and N. Shadbolt, "Agent-Based Semantic Web Services," in *Proceedings of the 12th International Conference on World Wide Web*, 2003, pp. 710–717.
- [6] S. Christley, X. Xiang, and G. Made, "AN ONTOLOGY FOR AGENT-BASED MODELING AND SIMULATION," 2004.
- [7] R. N. Giere, "An agent-based conception of models and scientific representation," *Synthese*, vol. 172, no. 2, p. 269, 2009.
- [8] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities," *Sci. Am.*, vol. 284, no. 5, pp. 1–5, 2001.
- [9] C. Bizer, T. Heath, and T. Berners-Lee, "Linked Data: The Story so Far," *Int. J. Semant. Web Inf. Syst.*, vol. 5, pp. 1–22, Jul. 2009.
- [10] M. Uschold, M. Gruninger, M. Uschold, and M. Gruninger, "Ontologies: Principles, Methods and Applications," no. February, 1996.
- [11] F. Van Harmelen and V. U. Amsterdam, "Handbook on Ontologies," *Handb. Ontol.*, no. December, 2009.
- [12] S. Parsons, *A Semantic Web Primer, Second Edition by Antoniou Grigoris and Harmelen Frank van, MIT Press, 288 pp., \$42.00*, vol. 24, no. 4. 2009.
- [13] L. Moreau, P. Groth, J. Cheney, T. Lebo, and S. Miles, "The rationale of PROV," *J. Web Semant.*, vol. 35, pp. 235–257, 2015.
- [14] R. G. Kyle, C. R. Ludka, X. Sim, and S. L. Court, "Proceedings of the 2000 Winter Simulation Conference J. A. Joines, R. R. Barton," *Simulation*, pp. 1347–1350, 2000.

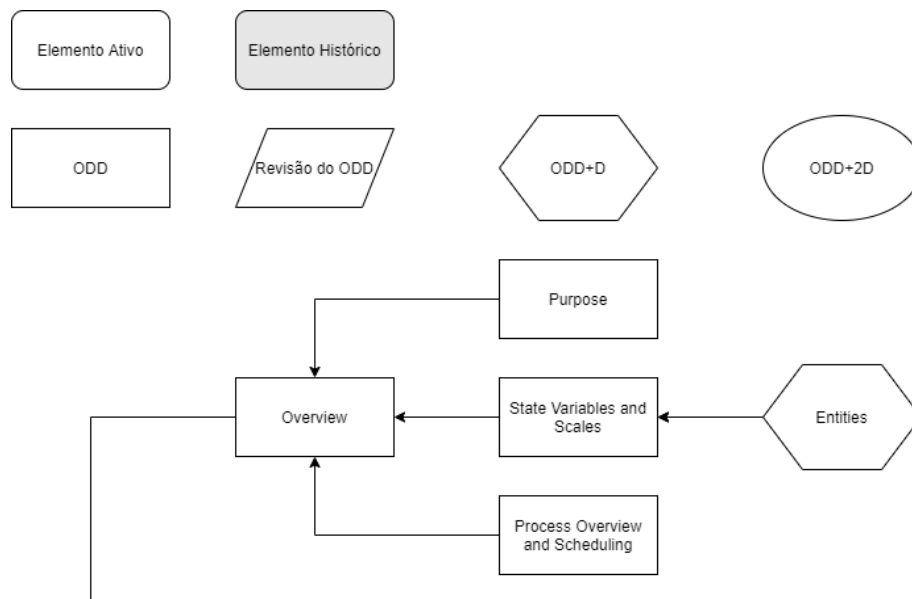
- [15] A. J. J. Heppenstall, A. Crooks, N. Malleson, E. Manley, J. Ge, and M. Batty, "Agent-based Models for Geographical Systems: A Review," 2019.
- [16] A. Crooks, N. Malleson, E. Manley, and A. J. Heppenstall, *Agent-based Modelling and Geographical Information Systems: A Practical Primer*. 2019.
- [17] A. J. J. Heppenstall, A. T. Crooks, L. M. See, and M. Batty, *Agent-based models of geographical systems*. 2012.
- [18] C. M. MacAI and M. J. North, "Tutorial on agent-based modelling and simulation," *J. Simul.*, 2010.
- [19] J. M. Epstein and R. Axtell, *Growing artificial societies: Social science from the bottom up*. Cambridge, MA, US: The MIT Press, 1996.
- [20] G. Bernardini, M. D'Orazio, E. Quagliarini, and L. Spalazzi, "An agent-based model for earthquake pedestrians' evacuation simulation in Urban scenarios," *Transp. Res. Procedia*, vol. 2, pp. 255–263, 2014.
- [21] A. Crooks, C. Castle, and M. Batty, "Key challenges in agent-based modelling for geo-spatial simulation," *Comput. Environ. Urban Syst.*, vol. 32, no. 6, pp. 417–430, 2008.
- [22] A. J. J. Heppenstall, A. T. Crooks, L. M. See, and M. Batty, *Agent-based models of geographical systems*, no. June. 2012.
- [23] F. Bousquet, G. Trébuil, and B. Hardy, *Companion Modeling and Multi-Agent Systems for Integrated Natural Resource Management in Asia*. 2005.
- [24] E. Xydas, C. Marmaras, and L. M. Cipcigan, "A multi-agent based scheduling algorithm for adaptive electric vehicles charging," *Appl. Energy*, vol. 177, pp. 354–365, 2016.
- [25] E. Silverman, J. Bijak, and J. Hilton, "Model-Based Demography in Practice: I," 2018, pp. 189–210.
- [26] Grimm V. *et al.*, "Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology," *Science (80-.)*, vol. 310, no. 5750, pp. 987–991, 2005.
- [27] V. Grimm, U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback, "The ODD protocol: A review and first update," *Ecol. Modell.*, vol. 221, no. 23, pp. 2760–2768, 2010.
- [28] B. Müller, F. Bohn, G. Dreßler, J. Groeneveld, C. Klassert, and J. Schulze, "0. ODD_revision2_complete_postprint," vol. 48, no. 0, pp. 37–48, 2013.
- [29] C. Wise, N. J. Car, R. Fraser, and G. Squire, "Standard provenance reporting and scientific software management in virtual laboratories," *Proc. - 21st Int. Congr. Model. Simulation, MODSIM 2015*, pp. 634–640, 2015.
- [30] C. Grueau and J. Araújo, *Towards a domain specific modeling language for agent-based models in land use science*. 2013.

- [31] K. Czarnecki and S. Helsen, "Feature-based survey of model transformation approaches," *IBM Syst. J.*, vol. 45, no. 3, pp. 621–645, 2006.
- [32] M. A. Musen, "The protégé project: a look back and a look forward," *{AI} Matters*, vol. 1, no. 4, pp. 4–12, 2015.
- [33] A. Bertolino, "Software Testing Forever: Old and New Processes and Techniques for Validating Today's Applications," in *PROFES*, 2008.
- [34] G. Guizzardi, "On ontology, ontologies, conceptualizations, modeling languages, and (meta)models," *Front. Artif. Intell. Appl.*, vol. 155, pp. 18–39, 2007.
- [35] J. Brank, M. Grobelnik, and D. Mladenić, "A survey of ontology evaluation techniques," 2009.

Apêndice I

Mapeamento ODD+2D parte 1

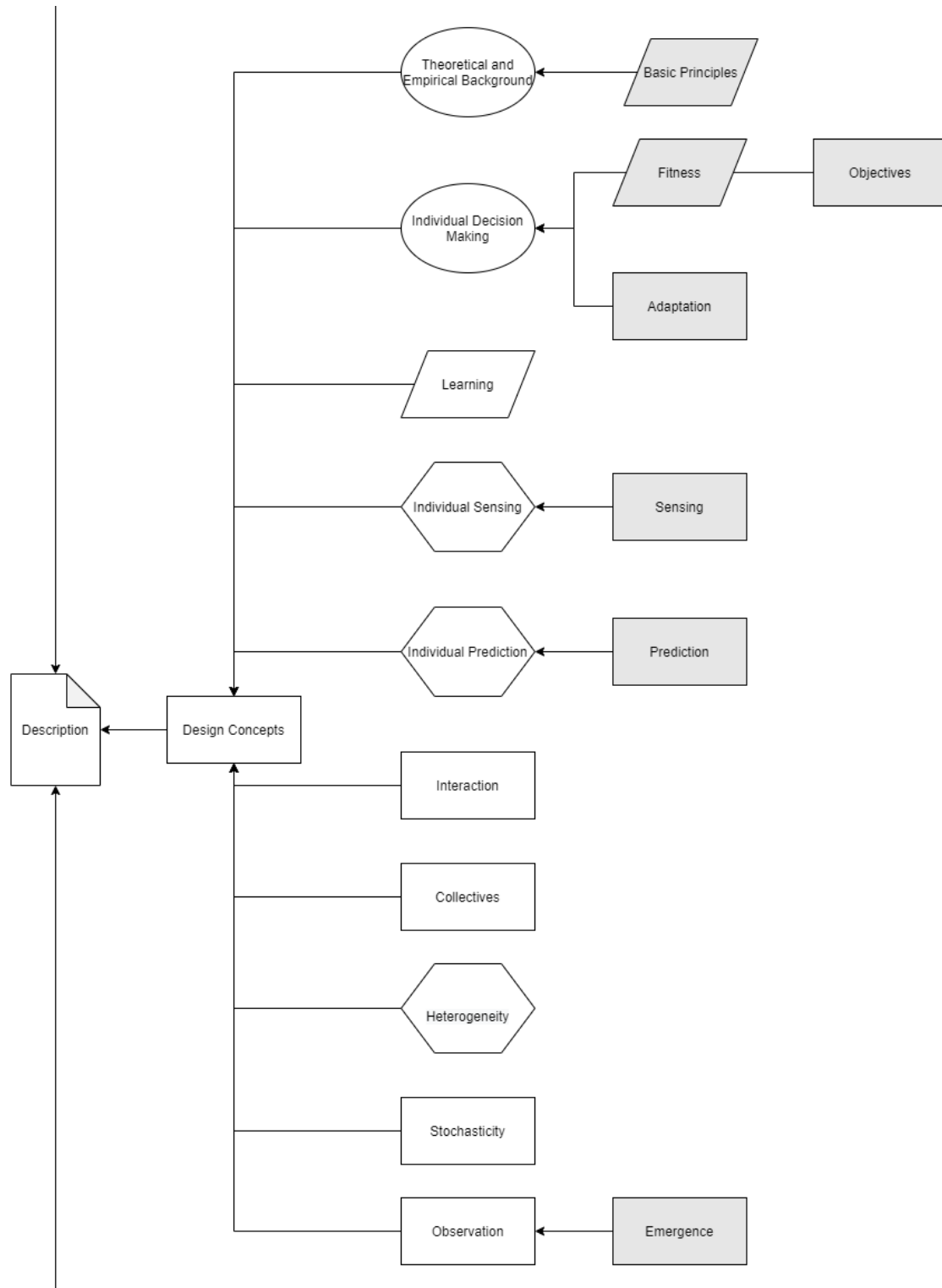
Primeira parte do mapeamento dos conceitos entre as versões do protocolo ODD+2D, contendo também a legenda para os outros apêndices: O retângulo representa a primeira versão do protocolo ODD, o paralelograma representa a revisão do protolo ODD, o hexágono representa o protocolo na versão ODD+2D, e a elipse representa a versão protocolo ODD+2D, os elementos em branco são os elementos que continuam a existir no protocolo ODD+2D, e os elementos em cinza são os que tiveram seus conceitos transferidos para outro elemento.



Apêndice II

Mapeamento ODD+2D parte 2

Segunda parte do mapeamento dos conceitos entre as versões do protocolo ODD+2D.



Apêndice III

Mapeamento ODD+2D parte 3

Terceira parte do mapeamento dos conceitos entre as versões do protocolo ODD+2D.

