Chapter 1 – Introduction

* Importance of the information systems, and their capability in word recognition. How to represent information.
* How to structure information. (Talk about controlled vocabularies).
* Ontologies as a structured representation of information.
* Domain ontologies as form of structure information in a specific domain.
* Maintaining and updating an ontology (ontology learning)
* From unstructured information to concepts and relations.
* What are concepts? Difficulties/challenges in capture/recognize concepts. How to capture / recognize concepts.
* What are relations? How to represent meaning. How to measure it. How recognize relations in words. Methods of quantification of relations of words?

Chapter 1.1 – Motivation

* No existence of a pure approach to quantify relations discovered from unstructured information in documents, without help of an ontology.
* What can be done to measure a relation and find its meaning?
* This document presents an approach to help discover relations in unstructured information in documents, knowing that there are no real methods to help measure a relation between two or more concepts.

Research Question:

How to quantify semantic relations between concepts in a domain ontology, using external sources of non-structured information.

Hypothesis:

Semantic relations between concepts from a domain ontology, can be quantified by applying data mining techniques for pattern extraction into non-structured sources of information.

Research questions:

* Having a set of documents with unstructured information, how could meaning be discovered, in the way of relations between its concepts?
* How to discover the domain of a set of words?

Chapter 1.2 – Goals

Present the way that I will propose solutions to research questions.

* How to address the problems?
* What techniques to use?
* Why are these techniques used to solve the problems, and not others?
* Develop a system, proof of concept, to present the results to domain experts.

Chapter 1.3 – Document Structure

Chapter 2 – Literature Review (State of the art)

Chapter 2.1 – Controlled Vocabularies (What are they? What do they represent?)

* What forms of representation of information exist?
* Ontologies (Definition, Construction, relations, concepts)

What is an ontology? What is it utility? How to construct one? Languages to represent it.

* Relations (meaning)
* Concepts
* Ontology learning
* E-cognos (new way of ontology creation).
* Application domain. (Practical cases where association rules are used)

Chapter 2.2 – Pattern Extraction

* Data mining. (What is DM? Techniques used today?)
* Association Rules (Definition, Rules)
  + Algorithms to discover [ECLAT, APRIORI, FP-GROWTH]
  + Weaknesses/Strengths between them
  + Why FP-Growth?
* Application domain. (Pratical cases where association rules are used)

Chapter 3 – Secção teórica (modelo/abordagem proposta) ???

Chapter 4 – Metodologia de desenvolvimento

- Method proposal to address the question.

- Techniques used to reach the goals.

- Technologies used

Chapter 5 – System Design and Implementation

- Relations discovered

- New concepts discovered

- Front end

Chapter 6 – Result Arguing and Analysis. Use cases

Chapter 7 – Conclusão e perspectivas futuras