Universiti Teknologi MARA

Semantic Annotation of Patent for Intellectual Property Document

Zeti Norasila Binti Zainal

Thesis submitted in fulfillment of the requirements for Bachelor of Computer Science (Hons) Faculty of Computer and Mathematical Sciences

July 2013

ACKNOWLEDGEMENT

ABSTRACT

The purpose of this project was to select and used one of the semantic annotation tools. The problem of lack of richer metadata from the patent document itself makes this method is appropriate to use. Semantic annotation tools can be divided into three which are manual annotation tool, semi-automatic annotation tool and automatic annotation tool. For the project, automatic semantic annotation tool is decided to use called AutoMeta. Therefore, from the tool the result of retrieving information is count as successful based on the calculation using the precision value (P). This value is counted for a reason to measure the performance of the search engine and the effectiveness of the retrieval process. Moreover, this project also shows how semantic search can save much time compared to using traditional search.

TABLE OF CONTENTS

CONTENTS   **PAGE**

CHAPTER 1

INTRODUCTION

This chapter provides the basic idea of this project. The topics that will be covered in these chapters are background study, problem statements, objectives, research scopes, significances and expected outcome of semantic annotation of patent for intellectual property document.

1.1 Background of Study

Nowadays, the internet becomes the important need for human. The internet gave more advantage for the users such as faster communication, social networking, online services and information resources. Information resource is the major advantage in the internet area because it help users can easily find many information based on their needs. Search engine likes Google and Yahoo are one of services on the internet. Millions of people used web search engine to help them to search their needs in the internet. However, don't people known that from using a web search engine it actually automatically engage them in the information retrieval area. This topic will explained more about the Information retrieval (IR) and Intellectual Property (IP) including with the history, its definition and many more.

1.1.1 Definition of Information Retrieval

Information Retrieval (IR) is a process of finding material usually unstructured data such as text that satisfy the information needs from the database (Christopher, Prabhakar & Hinrich, 2009). Bayes (2010) stated that area of IR had grown well. Therefore, there were many domains got involved and the popular domain is the domain of Intellectual Property.

1