

Cultural Heritage Knowledge Discovery: An Exploratory Study of the Sarawak Gazette

M.O.Rosita, R.Fatihah, K.M.Nazri, Alvin W.Yeo and Daniel Y.W.Tan

Department of Information Systems,
Faculty of Computer Science & Information Technology,
Universiti Malaysia Sarawak (UNIMAS),
94300 Kota Samarahan, Sarawak

{morosita, rfatihah, kmnazri, alvin}@fit.unimas.my
{leonard_tan84@yahoo.com}

Abstract. Sarawak Gazette has been used as a source of reference in social science research and is considered as one of the important repository of Sarawak's history, government and politics, people and their way of life, landscape, flora and fauna. In order to preserve this valuable source of information, the documents have been converted into digitized format. With the digitized format, further work can be done in order to allow not only text searches but also semantic searches. The objectives of this study are to identify structure features of the Sarawak Gazette and to find out whether the data can be represented in the form of an ontology. This is done by applying a text mining approach that is used to extract knowledge from the digital archives and to come out with a knowledge representation in the form of an ontology. This paper reports the preliminary findings of the project as well as the goals and challenges.

Keywords: ontology, semantic, historical archives, digital archives, cultural heritage, and knowledge discovery.

1 Introduction

In this paper, we take the definition of *cultural heritage* as national heritage or heritage is the legacy of physical artifacts and intangible attributes of a group or society that are inherited from past generations, maintained in the present and bestowed for the benefit of future generations [5]. According to UNESCO [6], the term "cultural heritage" covers several main categories of heritage which is tangible cultural heritage which means movable cultural heritage (paintings, sculptures, coins, manuscripts, etc.), immovable cultural heritage (monuments, archaeological sites, and so on), underwater cultural heritage (shipwrecks, underwater ruins and cities and so on) and intangible cultural heritage (oral traditions, performing arts, rituals, and so on). Cultural heritage also include natural heritage (natural sites with cultural aspects such as cultural landscapes, physical, biological or geological