i

A SCHEMA DRIVEN MAPPING:
MANAGING XML IN OBJECT-RELATIONAL DATABASE

A PROJECT REPORT PRESENTED BY

PIRITHAYINEE KANAGASABAPATHY

to the Board of Study in Statistics & Computer Science of the

POSTGRADUATE INSTITUTE OF SCIENCE

in partial fulfilment of the requirement
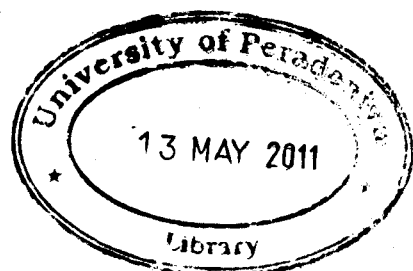*for the award of the degree of*

MASTER OF SCIENCE IN COMPUTER SCIENCE

of the

UNIVERSITY OF PERADENIYA

SRI LANKA

2010

A SCHEMA DRIVEN MAPPING:
MANAGING XML IN OBJECT-RELATIONAL DATABASE

P. Kanagasabapathy

Post Graduate Institute of Science

University of Peradeniya

Peradeniya

Sri Lanka

A new language namely XML (eXtensible Markup Language) has become ubiquitous and provides a crucial format for representing information on the World Wide Web. It is essential to determine techniques for managing XML documents. Accordingly, the relationship between XML and database management systems has become important. As such, it has become possible to store XML data in object-relational databases. Towards this, most of the prevailing techniques often exploit the stored XML data, usually expressed in Document Type Definition (DTD). However, the DTD became inadequate to represent today's complex applications and the need to use XML schema becomes more important.

In general XML schema [1] is used to describe the structure and the content of XML data. Using XML schema to store compliant data in database systems provides many benefits. However, the fidelity of the XML documents cannot be forfeited. Thus, the primary problem facing database developers is how XML schema can be mapped to object-relational database with no loss of schema semantics.

The project report describes flexible mapping of XML schema to object-relational database schema, outlines techniques for storing the valid XML data into relations of the resulting schema. The key objective is to exploit object-oriented features in XML schema. It also focuses on the modeling capabilities of object-relational databases to preserve the structure and the semantics of the data. This attempt is evaluated with the existing work and shows the significance of obtaining semantics in the mapping process.