Abstract

The growth of data available on the Internet and the improvement of ways to handle them consist of an important issue while designing a data model. In this context, XML provides the necessary formalism to establish a standard to represent and exchange data. Since the technologies of data warehouse are often used for data analysis, it is necessary to define a cube model data to XML. However, data representation in XML may generate syntactic, semantic and structural heterogeneity problems on XML documents, which are not considered by related approaches. To solve these problems, it is required the definition of a data schema. This paper proposes a metamodel to specify XML document cubes, based on relationships between elements and XML documents. This approach solves the XML data heterogeneity problems by taking advantages of data schema definition and relationships defined by XLink. The methodology used provides formal rules to define the concepts proposed. Following this formalism is then instantiated using XML Schema and XLink. It also presents a case study in the medical field and a comparison with XBRL Dimensions and a financial and multidimensional data model which uses XLink.

Keywords

XLDM, XML, XLink, XBRL, Multidimensional Data Metamodel.