Abstract

In the latest years many different models for semi structured data have been proposed; most of them, however, are too specific to allow immediate comparison with other models, and do not easily support incremental model design. A number of features which can be considered interesting for a semi structured data model are listed. E.g. we review the more widely used models in Conceptual Modeling for Information Systems (Entity-Relationship and UML), and argue that they do not support effectively modeling of semi structured data. As a consequence, structured and semi structured data cannot be treated in an integrated, holistic way during requirements specification. Research about similarity between semi structured documents (particularly for XML documents) has produced many works in the areas of Database Systems, Artificial Intelligence and Data Mining. In this work introducing a brief survey about it. At first introducing some basic properties. After that, some works are reviewed, highlighting their particularities and general approach. Concluding with a comparison of these works, analyzing their benefits and problems.

References
- Xiaoying Wu; Tok Wang Ling; Meng Li Lee; Dobbie, G.; "Designing semi structured databases using ORA-SS model"; In Proc. of the 2nd International Conference on Web
- Anirban Sarkar, "Conceptual Level Design of Semi-structured Database System:


Index Terms

Computer Science Database

Keywords

Semi-structured Data  XML  XSD  Conceptual Modeling  Semi-structured Data Modeling  XML Modeling