

XML is a popular format for data representation. As the amount of data represented in XML grows, it is necessary to concentrate on the process of modeling XML schemes of the XML representations. However, modeling the XML schemes on the level of XML schema languages, such as XML Schema, has some drawbacks. A natural idea to improve this situation is to model the XML schemes first on the conceptual level. It is motivated by the world of relational databases where we also start modeling the data first on the conceptual level. In this thesis we focus on conceptual modeling for XML. We start with a motivating example to point out to several problems that can arise when using only XML schema languages for modeling XML schemes. We discuss how modeling the data first on the conceptual level can help. We also show that conceptual modeling for XML has some specifics that should be taken into account by a conceptual model for XML. Mainly, we show that it is necessary to separate the conceptual modeling process to two parts. First, we need to model the data independently of its representation in XML. Second, we need to model how the data is represented in different types of XML documents. In the next step, we analyze in detail existing approaches to conceptual modeling for XML and show their limitations. In the main part of the thesis, we introduce our own conceptual model for XML called XSEM that extends the Entity-Relationship model and takes into account the specifics identified in the previous parts of the thesis. The required separation of the conceptual modeling process is supported by two subparts of XSEM called XSEM-ER and XSEM-H. We describe XSEM in a formal way and support the description with examples. Further, we describe in detail how to translate conceptual schemes in XSEM to a representation in the XML Schema language. We conclude with possible applications of the proposed model and our current and future work in the area.