Chapter III
XML and LDAP Integration: Issues and Trends

Vassiliki Koutsonikola
Aristotle University of Thessaloniki, Greece

Athena Vakali
Aristotle University of Thessaloniki, Greece

ABSTRACT

Nowadays, XML has become the standard for representing and exchanging data over the Web and several approaches have been proposed for efficiently managing, storing, querying and representing XML data originating from diverse and often heterogeneous sources. The Lightweight Directory Access Protocol is a promising technology for XML data storage and retrieval since it facilitates access to information organized under a variety of frameworks and applications. As an open, vendor-neutral standard, LDAP provides an extendable architecture for centralized storage and management of information that needs to be available for today’s distributed systems and services. The similarities between XML and LDAP data representation have led to the idea of processing XML data within the LDAP framework. This chapter focuses on the topic of LDAP and XML integration with emphasis on the storage and retrieval approaches implemented so far. Moreover, the chapter includes an overview and survey of the theoretical background and the adopted practices as realized in the most popular and emerging frameworks which tune XML and LDAP.

INTRODUCTION

Extensible markup language (XML) has rapidly emerged as the dominant standard for representing and exchanging data over the Web. As more and more enterprises take advantage of the connectivity offered by the Internet to exchange information within and across their boundaries, XML constitutes a simple yet extensible, platform-neutral, data representation standard.
XML’s simplicity and open nature has met Web data exchange requirements and nowadays the increasing amount of XML data on the Web poses the need for revising effective XML data management policies.

Among data management issues, storage and querying techniques are of particular importance, since the performance of an XML-based information system relies on them. Several approaches for storing XML data have been proposed in research literature and have been designed for various commercial tools. Apart from supporting the expected user-defined types of queries, an XML data storage system must also meet specific requirements such as:

- Supporting the management of XML schemas to define the data and the validation structures in order to map data to the pre-defined schemas;
- Providing mechanisms that will support effective indexing and disk space allocation methods in terms of the physical XML database storage;
- Integrating tools for content management operations (add, delete, modify) on the data organized under a specific storage framework;
- Supporting backup, replication and recovery mechanisms together with query optimization techniques based on indexing and other access paths configurations.

Although XML documents include text only and can easily be stored in files, it is worthwhile to “turn” to data management systems (e.g. databases) for their storage and retrieval, in order to apply advanced data management techniques. Figure 1 depicts the architecture of an XML-based storage and retrieval system. As depicted in this figure, appropriate storage policies are applied to the XML documents in order to store them in an XML data storage system while user applications or other application interfaces must use specific policies to retrieve the XML data. When integrating it with a database, the XML document structure has to be mapped to the database schema, which is required by every database management system. The structure of XML documents

![Figure 1. XML storage and retrieval](image-url)