

403431 Information Retrieval, Hypermedia & the Web

[3 cr., 3, 0] The objective of the course is that students will understand the fundamental techniques for document management and retrieval. Databases are not the only means for the storage, and subsequent retrieval of information, -- in fact databases only hold the subset of information known as "structured data". Although this constitutes the majority of data that drives the operational processes of an enterprise, it is actually the minority of information that is found in an enterprise. Documents and hypermedia are also information repositories, often referred to as semi-structured data, and forming the backbone of Digital Libraries. A Digital library is the combination of a digital repository, metadata describing objects in the repository, a set of users and systems that offer a variety of services (e.g. capture, indexing, cataloging, search, browsing, retrieval, delivery, archiving, and preservation). Work has gone on for at least a decade on how to manage and find electronic documents, and how to structure and navigate hypertexts. Work has been going on for centuries on how to manage and catalogue libraries. The WWW, as a global document repository and a distributed hypermedia, makes this area of information management more important than ever. A customer or another business finding my businesses web pages is a matter of my business's survival in e-Commerce land