

Managing Unstructured Information in Organizations: Applications of Document and Content Management Technologies

Chia-an Chao

Organizational Department, School of Business, Indiana State University

Terre Haute, IN 47809

bechao@isugw.indstate.edu

Keywords: unstructured information, document management, content management, taxonomy.

Submission Category: research in progress

Despite the proliferation of information technologies, many organizations are still struggling to find better ways to manage unstructured information. According to Gartner Group estimates, 80% of organizational information is unstructured. Unstructured information ranges from traditional text-based paper and electronic documents (such as invoices, insurance claim forms, loan applications, contracts, technical manual, and product descriptions), to non-text-based content (such as engineering schematics, patient EKG diagrams and x-rays), to e-mails, web pages, and source code. The wide variety of documents and content makes classification, indexing, and storage a challenge; it makes retrieval by different user groups difficult also.

Adding to the challenge is the dynamic nature of document and content management. Information is integral to business processes and rarely remains static. Managing the entire life cycle of information (which includes its creation/capture, approval, distribution, and deletion/archival), along with version control, security, and collaboration to facilitate multiple authorship poses additional challenges to organizations. A third area of challenge is to ensure document and content management comply with corporate governance standards and federal regulations. Given the sheer volume and variety of documents and content, their dynamic nature, and the need to comply with various standards and regulations, it may not come as a surprise that document and content management rank third in organizational operating costs, behind payroll and inventory.

This study focuses on effective management of unstructured data in two key industries: health care and financial services. Among findings of this study will be the taxonomy of unstructured data in these industries and technologies applied to the effective management of diverse documents and content.

The study of document and content management is a timely subject. A recent survey of 150 IT executives conducted by the Aberdeen Group on their purchasing plans revealed that content/document management systems are on the top of their purchase list, despite a minimal increase in IT budget. Given the interest in solutions for managing unstructured information to reduce cost and improve efficiency, a more thorough understanding of the complexity of unstructured information and technologies available for managing documents and content is an important subject for end-user information technology practitioners and researchers.