



GUEST EDITORIAL

Characterizing the evolving research on enterprise content management

Pasi Tyrväinen¹,
Tero Päivärinta²,
Airi Salminen^{3,5} and
Juhani Iivari⁴

¹University of Jyväskylä, Department of Computer Science and Information Systems, PO Box 35, FIN-40014, Finland; ²Agder University College, Department of Information Systems, PO Box 422, 4604, Kristiansand, Norway; ³University of Toronto, Faculty of Information Studies, 140 St. George Street, Toronto, ON M5S 3G6, Canada; ⁴University of Oulu, Department of Information Processing Science, University of Oulu, PO Box 3000, FIN-90014, Finland

Abstract

Innovations in network technologies in the 1990's have provided new ways to store and organize information to be shared by people and various information systems. The term Enterprise Content Management (ECM) has been widely adopted by software product vendors and practitioners to refer to technologies used to manage the content of assets like documents, web sites, intranets, and extranets in organizational or inter-organizational contexts. Despite this practical interest ECM has received only little attention in the information systems research community. This editorial argues that ECM provides an important and complex subfield of Information Systems. It provides a framework to stimulate and guide future research, and outlines research issues specific to the field of ECM.

European Journal of Information Systems (2006) 15, 627–634.

doi:10.1057/palgrave.ejis.3000648

Keywords: Enterprise content management, information-system, research framework, review, IRM, document management

Introduction

Innovations in network technologies in the 1990s have provided new ways to store and organize information to be shared by people and various information systems. Before, documents, files, records, and databases were the most common means to organize information (Sprague 1995). After the adoption of Internet technologies, organizations experienced a fast, partly uncontrolled growth of information assets in Web sites, intranets, and extranets. Natural language text, often augmented with graphics, pictures and other audio-visual material, forms a great portion of these information assets (Tyrväinen *et al.*, 2005). Practitioners, consultants, and professional communities have, since the late 1990s, actively discussed the problems related to content management in organizations and a great number of content management tools have been made available. The term *Enterprise Content Management* (ECM) has been widely adopted by software product vendors and practitioners to refer to technologies used to manage the content of assets like documents, web sites, intranets, and extranets in organizational or inter-organizational contexts. Although the term is extensively used, there is still a considerable confusion about the meaning of the concept (Smith & McKeen, 2003).

Compared to the extent of practical interest in the phenomenon, ECM has received only little attention in the information systems (IS) research community until lately (Smith & McKeen, 2003, Päivärinta & Munkvold, 2005), although there are long research traditions in areas closely related to ECM. These include records management, information resource management, and electronic document management (Sprague, 1995). Research on

⁵ On leave from University of Jyväskylä, Department of Computer Science and Information Systems, Finland

electronic document management in particular provides a basis to extend the investigation from document assets to the variety of information assets in contemporary organizations.

Even though, among practitioners, the term ECM usually refers to systems and technologies, ECM challenges in organizations are only partially technological. Development of new content management solutions in organizations requires new content production strategies and changes in business processes and work practices (Salminen *et al.*, 2006). Therefore, ECM provides an important and complex subfield of IS for the computing disciplines. Computer Science and Software Engineering may well produce various technological innovations for content management, but they tend to neglect the content aspect and the organizational context of ECM. As a result of the complexity of ECM, we attempt first to provide a framework to stimulate and guide future research. Then we outline research issues specific to the field of ECM, and introduce the two articles of the special issue.

A framework for ECM research

ECM forms a rich research phenomenon and can be viewed from four perspectives (Figure 1): content, technology, enterprise and process.

The research questions of the content perspective concern identification of content items, their semantics, structuring, and organization, as well as the creation and use of content both by human users and IS. Within the technology perspective, the research concerns the development of hardware, software, and standards for content management in organizational context. The enterprise provides the context where the content of various information assets should be managed. The enterprise perspective considers organizational, social, legal, and business issues of content management. Finally, research from the process perspective considers the development and deployment of new content management solutions in enterprises. The following sections introduce these perspectives in more detail, focusing especially on the content perspective.

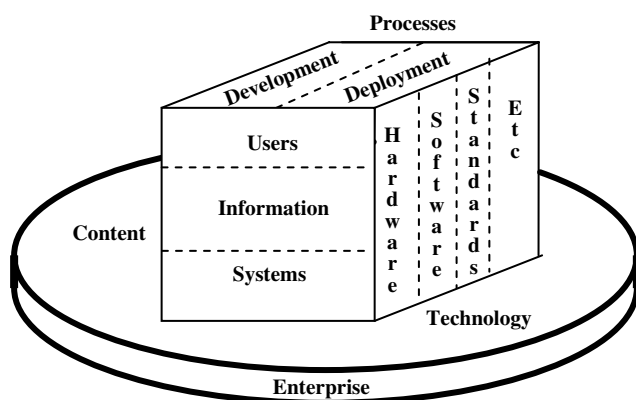


Figure 1 A framework for ECM research.

Content perspective

In any piece of ECM research, the content perspective is involved in some way. The word 'content' has a number of meanings, but essentially it always refers to something contained in an entity. Compared to the use of the terms 'data' or 'information', content clearly is associated with a container. We talk about the content of a document, content of a web site, or content of the Internet, among others. Content is often opposed to some other aspect of the container, for example, structure or form or representation. In an XML document, for example, we can separate content, structure, and one or more external presentations. Content carries the most substantial semantics of the entity. This is expressed also in the Oxford English Dictionary (second edition 1989) definition:

2.a. *spec. (pl.)* The things contained or treated of in a writing or document; the various subdivisions of its subject-matter, 2.b. *table of contents (content)*: a summary of the matters contained in a book, in the order in which they occur, usually placed at the beginning of the book, 3.a. The sum or substance of what is contained in a document; tenor, purport, 4. The sum of qualities, notions, ideal elements given in or composing a conception; the substance or matter (of cognition, or art, etc.) as opposed to the *form*.

Within the content perspective we have identified three views:

- *Information view* deals with the semantics of the content, how the content is represented and made accessible to users. Accordingly, it includes issues such as clustering content of an information repository into meaningful content units, granularity of content, metadata, text transformations, content search and retrieval, and content representation.
- *User view* focuses on the relationship between content and users, including both users creating and maintaining content and users reading and using the content. It covers questions such as 'who are users?', 'how do they read, interpret and use content?', 'how relevant is the content for them?', and 'how could the content and the system be personalized to satisfy their expectations and needs?'
- *System view* focuses on various systems in which the content resides and is made accessible to the users, on the granularity of content used for processing and storage, on interoperation of systems within and across organizational borders, and on the standards and formats related to these.

Information view

The word 'information' is problematic, since it is used in a variety of meanings. In our framework, information refers to the idea embedded in or conveyed by data (Iivari, 2005). That idea may be a simple fact or a complex piece of thought (Iivari, 2005). Datum (data) is an arrangement of physical symbols according to some language to represent and communicate some idea

(cf. Sundgren, 1973). The physical symbols cover oral speech, written text, audio-visual material, and computer representations. (We note that all digital content is not necessarily informational, for example music, which still may need to be managed as well for organizational purposes, for example, in businesses distributing digital entertainment products. However, the information view appears less important in such domains of content management.)

The *information view* is interested in the semantics of the content, how it is structured, represented, and made accessible to users. The early content management solutions used either technically coarse content storage granularity with very large content units (files, documents) or high granularity with very small content units (data items in databases). The structured document approach introduced the use of schemas for describing the structure of documents and adding semantic information to the content (Salminen & Tompa, 1999). The use of SGML and later, use of XML have followed this approach.

Recent research on the Semantic Web illustrates the increasing interest in the information view. The research was triggered by the perceived lack of semantic information related to the Internet content. The *Semantic Web* is intended to be 'an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation' (Berners-Lee *et al.*, 2001). The well-defined meaning is added to the web by means of metadata. The *metadata* is information about resources either accessible or identifiable on the web. Ontologies are used to express semantic metadata. An *ontology* defines formally the concepts and their relationships in an application domain (Gruninger & Lee, 2002). Management of information assets can be remarkably facilitated by the use of metadata. The schemas of structured documents are examples of metadata intended to support the management of documents. For example, in the content management model of Salminen (2005), the content items are divided into two classes: primary content items and metadata content items.

Accordingly, constructions of content models, metadata models and related ontologies have gained special attention in the ECM-related literature. XML-based schemas have been suggested to model the needs of content exchange and publishing in a number of application domains, such as health care (HL7, hl7.org), for e-commerce (ebXML, ebxml.org) and digital rights management (ODRL, odr1.net; Iannella, 2001; Rosenblatt *et al.*, 2002). These reflect the need to standardize information structures and semantics of global communication with a common data model. This requires also harmonizing the organizational processes across the collaborating enterprises demonstrating the tight interconnection of content management with process development and service oriented architectures. Use of ontology to structure domain concepts serves as a technical means to link the XML schemas and metadata

models, especially to natural language processing. The specification of ontologies for knowledge management in organizations and organization networks is an active area of research (e.g., Dieng-Kuntz *et al.*, 2006). As a result of wide adoption of XML technologies research has evolved to report experiences from XML-based standards for specific sectors, application domains, and organizations (e.g., Murray-Rust & Rzepa, 2001; Salminen *et al.*, 2001; Luk *et al.*, 2003). Metadata solutions for content management in organizational contexts have also been investigated in such fields as Archival Science (e.g., Gilliland *et al.*, 2005), Records Management research (e.g., Tough & Moss, 2003), and Information Resource Management research (e.g., Paganelli *et al.*, 2006).

Another research stream has approached the information view by analyzing how varying structural and technical forms of information content and communicative substances identified with the content emerge in an intertwined manner, resulting in recognizable *genres* of organizational communication (Yates & Orlikowski, 1992). Focus on genres of digital documents has provided a basis for content analysis in organizations (Tyrväinen & Päiväranta, 1999, Karjalainen *et al.*, 2000, Honkaranta & Lyytikäinen, 2003). Lately the genre theory has been developed further (Kwasnik & Crowston 2005) and applied to the analysis of various aspects related to user view of ECM and user communities as well as to genre systems related to organizational processes and to genre repertoires used by organizations. In this sense, the idea of genres of digital content represents a conceptual bridge between the information view and several other views when chosen as the unit of analysis in the ECM research.

All in all, a major problem in content management is to model and manage both the content communicating about some reality outside the container, and information about the ways the content is clustered, structured and stored in the container. Like in the traditional information retrieval, the users cannot retrieve information unless they have some information about documents as information containers; in ECM the users of information need knowledge of the ways information is stored and organized in their containers. For the users it would be important, however, that the information were stored and organized in semantic terms whenever possible. So far few successes have been achieved through the use of natural language processing techniques, and research has largely focused on content tagging using, for example, XML schemas, which provide the mainstream way to model and structure content for the majority of the contemporary solutions.

User view

Although we assume that the author of any information unit attempts to communicate some fact, idea or thought, the information content of the content unit is highly dependent on the reader as illustrated by Langefors' (1966) infological equation: $INF_r = i_r(D(INF_s), K_r, t)$.

This maintains that the information INF_r acquired by the receiver/reader r by interpreting the data or content unit $D(INF_s)$, representing an idea INF_s , is a function i_r of the data $D(INF_s)$ to be interpreted, the pre-knowledge K_r of the receiver/reader and the time t available for the interpretation.

This is even more so in the case of ECM where a content unit may include rich textual data often augmented with graphics, pictures, and other audio-visual material. Therefore, when creating, storing, distributing, and making content accessible, it is extremely important to anticipate the potential users and user communities as readers of the content, how they might read or misread the content, what is the context where the content is retrieved and used, how relevant the content is for them, and how it could be personalized to satisfy users' expectations and needs.

In the Information Retrieval (IR) community, the user-oriented IR research stream has lately shown increasing interest in information seeking and retrieval in social and organizational contexts (Ingwersen & Järvelin, 2005). For example, Stenmark (2005) has studied user deployment of information retrieval techniques in an international manufacturing corporation. However, beyond the applications of the genre theory to user-oriented information retrieval (Kwasnik & Crowston, 2005) and the user-oriented stream of IR research, the user view has been largely neglected in the ECM literature. Hence, more research is recommended from the content perspective on how the users are taken into account in modern ECM solutions and related user-oriented applications such as corporate portals.

Systems view

The system view focuses on systems (containers) where content units may reside in and/or through which content is made accessible to users. Such systems include document management systems, specific information systems, web sites, portals, etc. The system perspective is closely associated with the technology perspective to be introduced below but differs from it in the sense that systems have content and users, and they operate in some enterprise context (Figure 1). Large-scale content management systems can often incorporate and integrate a number of technologies.

A varied body of literature touches the system view. Sprague (1995) reviews the functionality of document management systems in the 1990's, Mack *et al.* (2001) outline functionality of enterprise portals, and Grossniklaus & Norrie (2002) argue for their generic model for content management systems. Also the IEEE EDM Workshop 'Electronic Document Management in an Enterprise Computing Environment' related to the EDOC Conference focuses on systems view combining technology, processes, data management, and some content related issues. The problem is, however, the fragmentation of the systems-oriented literature. Each new type of system technology tends to introduce its own concepts and

vocabulary, making a functional comparison of ECM-related systems difficult.

Fewer studies report experiences from actual ECM system implementations in organizations. Among the rare exceptions, Spahni (2004) reports how uniform resource name (URN) technology was utilized in distributed Swiss government and Becking *et al.* (2005) describe functionality of a content management product originating in a system implementation of a publishing company. Owing to the scattered nature of the system-oriented literature and lack of cumulative reporting from systems implementation experiences, we can speculate that system architectures in organizations resulting from such market-oriented development may often seem rather accidental than well-thought.

Technology perspective

While the system view of content perspective focuses on ECM applications accessible to the users, the technology perspective focuses more on the base technologies such as the software and hardware used in ECM. There is a rich body of the literature addressing ECM from the technology perspective (Päivärinta & Munkvold, 2005). Typically, this literature focuses on one or more generic technological ideas related to ECM without addressing much the content perspective or the organizational context of technology utilization. For example, Ahmed *et al.* (2003) suggest an algorithm for video segmentation, indexing and framing, Di Iorio & Vitali (2005) describe generic functional features of a Wiki-based co-authoring solution, Lee & Goodwin (2006) develop tools and techniques for ontology management, and Watters & Zhang (2003) suggest and evaluate a generic algorithm for making web sites viewable through hand-held screens.

Although technology is without doubt significant as an enabler of ECM, we argue that the major focus of ECM research from the Information Systems perspective lies in systems rather in individual technologies. A system requires and integrates a number of technologies.

Enterprise perspective

The enterprise aspect considers organizational, social, and business issues of content management. ECM research from the enterprise perspective is very limited, consisting mostly of early conceptual and theoretical recommendations and a limited set of empirical studies. On the other hand, the enterprise perspective is often tightly intertwined with the process perspective, which takes the viewpoint of development and deployment processes.

The impact of communication on enterprise performance is well known (Allen & Cohen 1969) and much of ECM research including the enterprise perspective has focused on the role of ECM in the communication processes of an enterprise. Several researchers have proposed that analysing organizational communication provides a basis for information systems and software (see Ågerfalk *et al.*, 2006). For example, the enterprise

perspective has been approached through the above-mentioned genre theory of organizational communication (Tyrväinen & Päivärinta, 1999; Karjalainen *et al.*, 2000), which intertwines the organizational purposes, tasks and user communities of the content to the explanations of why particular content forms emerge.

Among the other conceptual ideas to capture the enterprise perspective, Grossniklaus & Norrie (2002) and Han & Pape (2002) outline solutions to combine workflow modelling with content management. Forsyth (2005) discusses the role of content management in general in increasing the sales effectiveness in the domain of medical marketing.

However, empirical research including the enterprise perspective of ECM implementations remains rare. Munkvold *et al.* (2003) present an in-depth case study on ECM problems and issues in an oil corporation. Päivärinta & Munkvold (2005) review 58 commercially reported success stories, outlining a set of organizational issues related to ECM. However, actual examples of how the enterprise aspect has been actually modelled and the models deployed in the resulting solutions remain rare in the field.

Process perspective

The process perspective distinguishes between two major process categories related to information systems such as ECM:

- Development including the processes of developing, implementing and maintaining ECM systems with related change management.
- Deployment processes including implementation of content life-cycle activities such as content creation and capture, content editing, review, approval, content indexing, classifying and linking, content distribution, publication and use, update, preservation, format transformation for long-term archival, and retention (Gupta *et al.*, 2001; Smith & McKeen, 2003; Päivärinta & Munkvold, 2005).

Research on the development process includes methodological recommendations for development of ECM solutions in organizations and studies reporting experiences of actual development initiatives. In general, a few methodologies and techniques for developing content and document management systems have been proposed and tried out in varying development initiatives (Päivärinta, 2001; Salminen, 2003). Although it remains a rarity to critically report experiences from development methodologies beyond descriptions of the actual methods, such research has begun to emerge for example, with regard to variants of the genre-based approach (Päivärinta, 2001; Honkaranta & Lyytikäinen, 2003, Honkaranta & Tyrväinen, 2003) and in connection to investment justification techniques in ECM (Hallikainen *et al.*, 2002).

Another stream of literature has studied, without methodological preconceptions, organizational development and change related to ECM. Smith & McKeen (2003) address the issue of forming an ECM strategy based on

data from a focus group of Canadian managers. Scott *et al.* (2004) study an evolutionary process of ECM development in a software company and Junco *et al.* (2005) sketch lessons learned from an ECM implementation process in a medical equipment manufacturer. Honkaranta *et al.* (2005) describe experiences gained from renewing content models in a Finnish public sector organization.

At the level of deployment, there exists even less research, perhaps due to the lack of reports from organizational implementations of ECM in general. However, case studies of experiences from ontology and metadata deployment have been reported (e.g., Dieng-Kuntz *et al.*, 2006; Jokinen *et al.*, 2005).

Specific ECM research themes

There exists also a set of special ECM research themes that deserve special attention. Architecture, infrastructure and standards can be viewed from the technical perspective, but require attention also from other perspectives. The content perspective is very relevant for content architectures represented in the form of standardized genre repertoires and the enterprise perspective is relevant for interoperable data models in content messaging protocols, as, for example, in the study of Chen (2003) on factors affecting the adoption and diffusion of XML and web services standards for e-business systems.

The significance of security of the ECM architecture and technology is accentuated since ECM may include sensitive information such as business strategies, business plans, customers, marketing, product specifications, financial situation, and R&D secrets. This content may be of strategic value to the enterprise so that it is vital that the content is not lost, that it is kept up-to-date and that it is not disclosed to unauthorized people. These issues touch especially the financial industry (Kwok & Chiu, 2004; Chiu & Hung, 2005). There are also generic issues related to digital rights protection and management systems and models for them, and these need to be considered with regard to certain ECM solutions (Cheung & Chiu, 2003; Cheung *et al.*, 2004; Luoma & Tyrväinen, 2004; Kienle *et al.*, forthcoming). ECM systems may also unintentionally violate intellectual property rights of other organizations (Kienle *et al.*, forthcoming).

Summary

The framework of Figure 1 emphasizes that ECM as a research phenomenon is multi-dimensional. A piece of research may address ECM from a number of perspectives. We claim, however, that any piece of ECM research should include the content perspective. Accordingly, Figure 1 takes the content perspective into the front. This distinguishes ECM from traditional IS. Iivari's (1989) framework of information systems, for example, takes systems into the center, identifying organizational (enterprise) perspective, information perspective, and technology perspective in information systems. The user perspective in Iivari (1989) is included in the enterprise perspective. The focus on content rather than systems is

important for two reasons. First, a part of content may reside in and be accessible through different systems and technologies. A second reason is that in many application domains, such as in health care or product information, content may be required to be stored for decades or centuries. So, the content should be able to live through generations of technologies, systems, users, and even surrounding organizations.

With regard to research approaches, ECM research tends to favour constructive research typical to design science (Hevner *et al.*, 2004), conceptual models, and qualitative empirical works. Evaluative works tend to be in the minority, with the exception of information retrieval studies having a strong tradition in laboratory experiments. Depending on the focus the unit of analysis may be an organizational unit, a content unit at some level of granularity, a genre or organizational communication, or a special issue of interest (such as digital rights management) to name a few. ECM shares a wide variety of reference disciplines with IS research in general, whereas the prominence of the information view suggests an especially strong focus on semantics and organizational communication as the key disciplines. Technology-oriented ECM research relates more to areas of Computer Science, such as grammars, algorithms, and cryptography.

Special section

The Call for Papers for this special issue was distributed in 2004, including the papers accepted to the ECM Mini-

track in HICSS during years 2003 to 2005 (Tyrväinen *et al.* 2003). Nine submissions were received. Two of these submissions were finally accepted.

The article of Sheepers 'A conceptual framework for the implementation of enterprise information portals in large organizations' suggests a model to aid user segmentation and personalization during the development process of enterprise portals, with preliminary evidence of usefulness of such ideas based on two case studies. It suggests the 4P model from marketing to inform system developers for segmentation, customization, personalization and promotion of enterprise information portals. Referring to Figure 1, the focus of the paper thus lies in the development process focusing especially on user (customer) and system views within the content perspective to facilitate adoption of new solutions.

The article of Nordheim and Päiväranta 'Implementing enterprise content management: From evolution through strategy to contradictions out-of-the-box' discusses holistic implementation of an ECM system in Statoil, a large Norwegian oil company. It applies the framework of organizational change proposed by Van de Ven & Poole (1995) to interpret the ECM implementation process in the target organization and suggests that ECM developers should consider various ideas of change that affect the process and perceived results from the development efforts. Referring to Figure 1, the major focus of this article thus lies in the development process from an enterprise perspective.

About the authors

Dr. Pasi Tyrväinen is Professor of Computer Science and Information Systems (Digital Media) at University of Jyväskylä. He received doctoral degree at Helsinki University of Technology in 1994. His previous affiliations include R&D management positions at Honeywell Industrial Control and Nokia Research Center. His research interests include enterprise content management, communication genres, DRM, and software business.

Dr. Tero Päiväranta is Associate Professor at Agder University College, Norway. His current research interests include e-democracy, enterprise content management, genre theory in ISD, and benefits management. His works appear in more than 20 IS conferences and journals, for example Information & Organization, Communications of the AIS, and Information Systems Journal.

Dr. Airi Salminen is Professor at the University of Jyväskylä, Finland. Currently, she serves as a visiting professor at the Faculty of Information Studies of the University of Toronto. She has headed several research projects on CM development efforts in major Finnish companies and public sector organizations. Her current research interests include ECM, structured documents, XML, e-government, and semantic web.

Dr. Juhani Iivari is a Professor in Information Systems at the University of Oulu, Finland, and the Scientific Head of the INFWEST Education Program in the area of information systems. His research interest lies in theoretical foundations of information systems, ISD methodologies and approaches, acceptance of IT applications, IS quality, and in the relationship between information systems and knowledge work.

References

ÅGERFALK PJ, GOLDKUHL G, FITZGERALD B and BANNON L (2006) Reflecting an action in language, organisations and information systems. *European Journal of Information Systems* 15(1), 4–8.

AHMED M, IMPEY R and KARMOUCH A (2003) Developing video services for mobile users. In *Proceedings of the 36th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.

- ALLEN TJ and COHEN SI (1969) Information flow in research and development laboratories. *Administrative Science Quarterly* **14**(1), 12–19.
- BECKING J, COURSE S, VAN ENK G, HANGYI HT, LAHAYE JJM, OCKELOEN D, PETERS R, ROSBERGEN H and VAN WENDEL DE JOOTE R (2005) MMBase: An open-source content management system. *IBM Systems Journal* **44**(2), 381–397.
- BERNERS-LEE T, HENDLER J and LASSILA O (2001) The Semantic Web. *Scientific American* **284**(5), 34–43.
- CHEN M (2003) Factors affecting the adoption and diffusion of XML and Web services standards for E-business systems. *Human-Computer Studies* **58**(3), 259–279.
- CHEUNG SC and CHIU DKW (2003) A watermarking infrastructure for enterprise document management. In *Proceedings of the 36th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- CHEUNG S, LEUNG H and WANG C (2004) A commutative encrypted protocol for the privacy protection of watermarks in digital contents. In *Proceedings of the 37th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- CHIU DKW and HUNG PCK (2005) Privacy and access control issues in financial enterprise content management. In *Proceedings of the 38th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- DIENG-KUNTZ R, MINIER D, RUZICKA M, CORBY F, CORBY O and ALAMARGUY L (2006) Building and using a medical ontology for knowledge management and cooperative work in a health care network. *Computers in Biology and Medicine* **36**(7–8), 871–892.
- DI IORIO A and VITALI F (2005) Web authoring – a closed case?. In *Proceedings of the 38th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- FORSYTH K (2005) Content management: A prerequisite to marketing and sales effectiveness. *International Journal of Medical Marketing* **4**(3), 228–234.
- GILLILAND A, ROUSSE N, LINDBERG L and EVANS J (2005) Towards a 21st century metadata infrastructure supporting the creation, preservation and use of trustworthy records: Developing the InterPARES 2 metadata schema registry. *Archival Science* **5**(1), 43–78.
- GROSSNIKLAKS M and NORRIE MC (2002) Information concepts for content management. In *Proceedings of the Third International Conference on Web Information Systems Engineering (Workshops)*. IEEE Computer Society, Los Alamitos.
- GRUNINGER M and LEE J (2002) Ontology applications and design. *Communications of the ACM* **45**(2), 39–41.
- GUPTA VK, GOVINDARAJAN S and JOHNSON T (2001) Overview of content management approaches and strategies. *Electronic Markets* **11**(4), 282–288.
- HALLIKAINEN P, KIVIJÄRVI H and NURMIMÄKI K (2002) Evaluating strategic IT investments: An assessment of investment alternatives for a web content management system. In *Proceedings of the 35th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- HAN X and PAPE U (2002) The integration of content management system and workflow management system – a conceptual model. In *Proceedings of International Conference on e-Business (ICEB 2002)* May 23–26 Beijing.
- HEVNER AR, MARCH ST, PARK J and RAM S (2004) Design science in information systems research. *MIS Quarterly* **28**(1), 75–105.
- HL7 (1999) HL7 reference information model. http://www.hl7.org/library/data-model/RIM/modelpage_non.htm (accessed September 26, 2006).
- HONKARANTA A and LYYTIKÄINEN V (2003) Operationalizing a genre-based method for content analysis: A case of a church. In *Proceedings of the 6th International Conference on Business Information Systems* (ABRAMOVICZ W and KLEIN G, Eds), pp. 108–116, Poznan University of Economics, Poznan.
- HONKARANTA A, SALMINEN A and PELTOLA T (2005) Challenges in the redesign of content management: Case of FCP. *International Journal of Cases on Electronic Commerce* **1**(1), 44–60.
- HONKARANTA A and TYRVÄINEN P (2003) Designing training in manufacturing organizations using the genre-based method. *Journal of Educational Technology and Society* **6**(4), 73–85.
- IANNELLA R (2001) Digital rights management (DRM) architectures. *D-Lib Magazine* **7**(6) <http://www.dlib.org/dlib/june01/iannella/06iannella.html> (accessed September 9 2006).
- IVARI J (1989) Levels of abstraction as a conceptual framework for an information system. In *Information Systems Concepts: An In-Depth Analysis* (FALKENBERG, ED and LINDGREEN, P, Eds), pp. 323–352, North-Holland, Amsterdam.
- IVARI J (2005) Is the future of the Scandinavian information systems development approaches passé. In *History of Nordic Computing* (BUBENKO J Jr, IMPAGLIAZZO J and SOLVBERG A, Eds), pp. 339–356, Springer, New York, NY.
- INGWERSEN P and JÄRVELIN K (2005) *The Turn. Integration of Information Seeking and Retrieval in Context*. Springer, Dordrecht, The Netherlands.
- JOJINEN K, BORGMAN J and SULONEN R (2005) Common data model for design document exchange in business to business networks. In *Proceedings of the 38th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- JUNCO N, BAILIE RA and LEDET D (2005) A case study of a content management system: Choosing and implementing a CMS. In *2005 IEEE International Professional Communication Conference Proceedings* pp. 343–345, IEEE, Los Alamitos.
- KARJALAINEN A, PÄIVÄRINTA T, TYRVÄINEN P and RAJALA J (2000) Genre-based metadata for enterprise document management. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. IEEE Computer Society, Los Alamitos CA.
- KIENLE H, GERMAN D, TILLEY S and MÜLLER H (forthcoming) Managing legal risks associated with intellectual property on the web. To appear in. *International Journal of Business Information Systems* **3**(1).
- KWASNIK BH and CROWSTON K (2005) Introduction to the special issue: Genres of digital documents. *Information Technology & People* **18**(2), 76–88.
- KWOK KHS and CHIU DKW (2004) A web services implementation framework for financial enterprise content management. In *Proceedings of the 38th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- LANGFORS B (1966) *Theoretical Analysis of Information Systems*. Studentlitteratur, Lund, Sweden.
- LEE J and GOODWIN R (2006) Ontology management for large-scale enterprise systems. *Electronic Commerce Research and Applications* **5**(1), 2–15.
- LUK RWP, T'SOU BKY, LAI TBY, KWONG OOO, CHIK FCY and CHEUNG LYL (2003) Bilingual legal document retrieval and management using XML. *Software Practice & Experience* **33**(1), 41–59.
- LUOMA E and TYRVÄINEN P (2004) Integrated domain model for digital rights management. In *E-Commerce and M-Commerce Technologies* (DEAN PC, Ed.), pp. 142–165, IRM Press, Hershey, PA, USA.
- MACK R, RAVIN Y and BYRD RJ (2001) Knowledge portals and the emerging digital knowledge workplace. *IBM Systems Journal* **40**(4), 925–955.
- MUNKVOLD BE, PÄIVÄRINTA T, HODNE AK and STANGELAND E (2003) Contemporary issues of enterprise content management: the case of Statoil. In *Proceedings of the 11th European Conference on Information Systems* (CIBORRA CU, MERCURIO R, DE MARCO M, MARTINEZ M and CARIGNANI A, Eds), CD-ROM, Naples, Italy.
- MURRAY-RUST P and RZEPA HS (2001) Chemical markup, XML, and the World Wide Web. 3. Toward a signed semantic chemical web of trust. *Journal of Chemical Information and Computer Sciences* **41**(5), 1124–1130.
- PAGANELLI F, PETTENATI MC and GIULI D (2006) A metadata-based approach for unstructured document management in organizations. *Information Resources Management Journal* **19**(1), 1–22.
- PÄIVÄRINTA T (2001) A genre-based approach to developing electronic document management in the organization. PhD Dissertation. *Jyväskylä Studies in Computing, Vol. 11*. University of Jyväskylä, Jyväskylä.
- PÄIVÄRINTA T and MUNKVOLD BE (2005) Enterprise content management: an integrated perspective on information management. In *Proceedings of the 38th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- ROSENBLATT B, TRIPPE B and MOONEY S (2002) *Digital Rights Management: Business and Technology*. M&T Books, New York.
- SALMINEN A (2003) Document analysis methods. In *Encyclopedia of Library and Information Science* (BERNIE CL, Ed.), pp. 916–927, Marcel Dekker, New York.

- SALMINEN A (2005) Building digital government by XML. In *Proceedings of the Thirty-Eighth Hawaii International Conference on System Sciences* (SPRAGUE RH Jr. Ed.), IEEE Computer Society, Los Alamitos, CA.
- SALMINEN A, LYYTIKÄINEN V, TIITINEN P and MUSTAJÄRVI O (2001) Experiences of SGML standardization: The case of the Finnish legislative documents. In *Proceedings of the Thirty-Fourth Hawaii International Conference on System Sciences* (SPRAGUE RH Jr. Ed.), IEEE Computer Society, Los Alamitos, CA.
- SALMINEN A, NURMEKSELA R, LEHTINEN A, LYYTIKÄINEN V and MUSTAJÄRVI O (2006) Content production strategies for e-Government. In *Encyclopedia of Digital Government* (ANTTIROIKO A-V and MÄLKÄ M, Eds), pp. 224–230, IDEA Group Publishing, Hershey, PA.
- SALMINEN A and TOMPA F (1999) Grammars++ for modelling information in text. *Information Systems* **24**(1), 1–24.
- SCOTT J, GLOBE A and SCHIFFNER K (2004) Jungles and gardens: the evolution of knowledge management at JD Edwards. *MIS Quarterly Executive* **3**(1), 37–52.
- SMITH HA and MCKEEN JD (2003) Developments in practice viii: enterprise content management. *Communications of AIS* **11**(33), 1–26.
- SPAHLI D (2004) Managing access to distributed resources. In *Proceedings of the 36th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- SPRAGUE RH (1995) Electronic document management: Challenges and opportunities for information systems manager. *MIS Quarterly* **19**(1), 29–49.
- STENMARK D (2005) Query expansion on a corporate intranet: Using LSI to increase precision in explorative search. In *Proceedings of the 38th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- SUNDGREN B (1973) *An Infological Approach to Data Bases*. Skriftserie utgiven av statistiska centralbyrån, Nummer 7, Statistiska Centralbyrån, Stockholm.
- TOUGH A and MOSS M (2003) Metadata, controlled vocabulary and directories: electronic document management and standards for records management. *Records Management Journal* **13**(1), 24–31.
- TYRVÄINEN P and PÄIVÄRINTA T (1999) On rethinking organizational document genres for electronic document management. In *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), IEEE Computer Society, Los Alamitos, CA.
- TYRVÄINEN P, SALMINEN A and PÄIVÄRINTA T (2003) Introduction to the enterprise content management minitrack. In *Proceedings of the 36th Hawaii International Conference on Systems Sciences* (SPRAGUE RH Jr. Ed.), IEEE Computer Society, Los Alamitos, CA.
- TYRVÄINEN P, KILPELÄINEN T and JÄRVENPÄÄ M (2005) Patterns and measures of digitalisation in business unit communication. *International Journal of Business Information Systems* **1**(1/2), 199–219.
- VAN DE VEN AH and POOLE MS (1995) Explaining development and change in organizations. *Academy of Management Review* **20**(3), 510–540.
- WATTERS C and ZHANG R (2003) PDA access to Internet content: focus on forms. In *Proceedings of the 36th Hawaii International Conference on System Sciences* (SPRAGUE RH Jr, Ed.), CD-ROM, IEEE, Los Alamitos.
- YATES J and ORLIKOWSKI WJ (1992) Genres of organizational communication: A structural approach to studying communication and media. *Academy of Management Review* **17**(2), 299–326.