7 Systems

KMS were defined in section 4.3 - “Knowledge management systems” on page 82. In the following, first the technological roots of KMS are reviewed (section 7.1). Then, the contents of KMS are analyzed along with their structure, the types of media used, a maturity model for knowledge elements and some aspects of quality of contents (section 7.2). The definition of KMS is detailed with the help of a review of KMS architectures that have been proposed in the literature or have been implemented as standard KMS platforms. Based on this analysis, an amalgamated architecture for a centralized KMS is presented. The architecture is discussed in detail with the help of a structured list of KMS functions that will be used in the empirical study (section 7.4). As an alternative to this ideal architecture for a centralized KMS, an architecture for a distributed or peer-to-peer KMS is presented (section 7.5). The development of tools and systems will be discussed in a structured way leading to a classification of KMS (section 7.6). Finally, the important integration layer is discussed in more detail, reflecting on meta-data and ontology management as well as the Semantic Web (section 7.7).

7.1 Technological roots

Figure B-47 uses the metaphor of a magnetic field produced by a coil to show the technological roots and influences that impact design and implementation of KMS. The term KMS plays the role of the coil, the magnetic center. Theoretical approaches that support deployment of KMS and related terms that show a different perspective on ICT support of an organization’s way of handling knowledge are shown to the right of the magnetic center. The main differences between KMS and their predecessors guiding the design of KMS are shown on the left side[440]. Both influences together provide the energy to integrate, (re-) interpret, (re-)arrange and (re-) combine ICT technologies that are the roots of KMS into a set of KMS-specific services that in turn are integrated into application systems, tools and platforms with a clear focus on the support of KM concepts and instruments.

The strong metaphor of a KMS, a system aiding the handling of knowledge in an organization, influences other ICT-related initiatives that can benefit from the ideas integrated with the help of KMS. Examples are the overall handling of electronic assets in an enterprise-wide content management, the integration of intelligent services for strategic enterprise management, the provision of access from any location in mobile information management, the specialized management of knowledge about employees, customers, projects, processes and products, the support of training and education by e-learning as well as the personal knowledge management of networked knowledge workers.

440. For an explanation see section 4.3.2 - “Definition” on page 86.

In the following, the most important ICT will be reviewed that form the technological roots of KMS[441]. Comprehensive KMS combine and integrate the functionality of several, if not all of these predecessors:

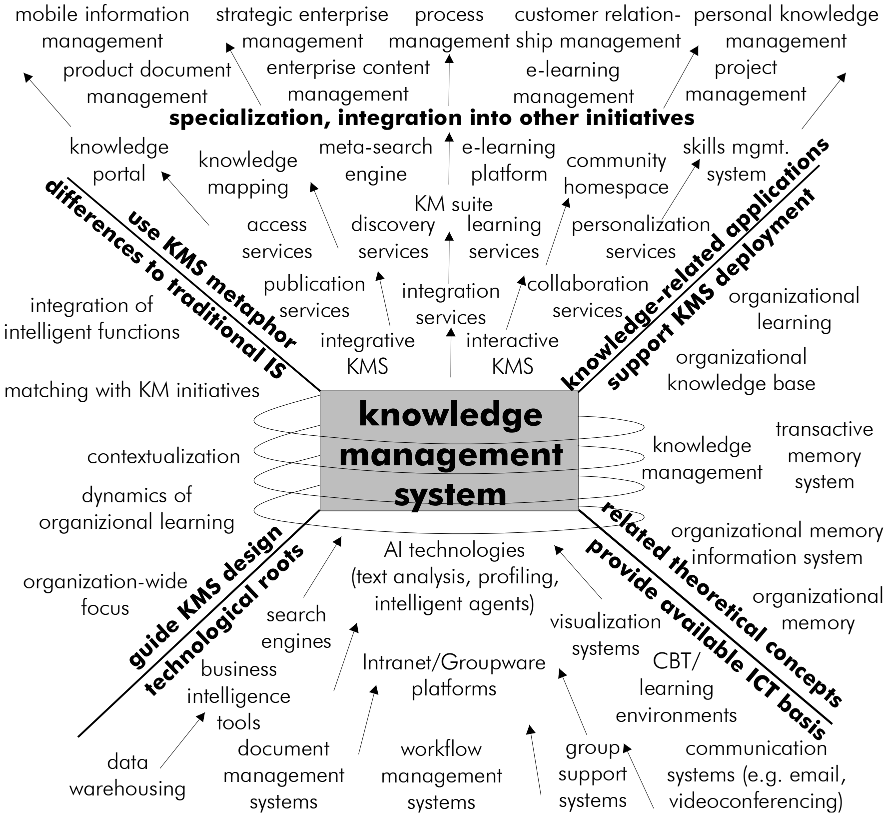

FIGURE B-47. Technological roots and influences of KMS

Document and content management. The term document management denotes the automated control of electronic documents, both individual and compound documents, through their entire life cycle within an organization, from initial creation to final archiving (Turban et al. 1999, 433f), i.e., creation, storage, organization, transmission, retrieval, manipulation, update and eventual disposition of documents (Sprague 1995, 32). A document management system (DMS) provides functions to store and archive documents, navigate and search documents, for versioning and to control access to documents. Additionally, many DMS support the pro-

441. See chapter 7 - "Systems" on page 273 for a detailed discussion of the various services, applications and specializations of KMS.

cess of imaging which turns paper-based documents into electronic ones and the classification of documents (Mertens et al. 1997, 128f, Thiesse/Bach 1999, 100ff).

A content management system (CMS) supports the organization of information and contents and the publication on the Web. Like DMS in the non-Web environment, CMS manage the whole Web publishing process, offer mechanisms for releasing new contents, support HTML generation with the help of templates, standard input and output screens and the separation of content and layout which provides for a standardized look & feel of the Web pages (Horn 1999, 165). As a consequence, participants who are not familiar with HTML can publish Web documents that fit into an organization's corporate (Web) identity. So-called Wikis and Weblogs are purpose-oriented CMS that are pre-structured, offer a subset of easy-to-use CMS functions and allow for simple (joint) editing, updating and linking of content within and between sites[442].

Workflow management. A workflow is the operative, technological counter-part of a business process and consists of activities related to one another which are triggered by external events and carried out by persons using resources such as documents, application software and data (Galler 1997, 7f). A workflow management system (WFMS) "defines, creates and manages the execution of workflows through the use of software, running on one or more workflow engines, which is able to interpret the process definition, interact with workflow participants and, where required, invoke the use of IT tools and applications" (WfMC 1999, 9, for examples for WFMS see Koch/Zielke 1996, 162ff). Most WFMS primarily support well-structured organizational processes. More recently, some WFMS also focus flexible workflows, so-called ad-hoc workflows (Galler 1997, 16f). An ad-hoc workflow is a sequence of tasks that cannot be standardized, but has to be designed spontaneously by participants (Koch/Zielke 1996, 30). WFMS functionality can be used in knowledge management, e.g., to support processes such as the publication or distribution of knowledge elements. Several KMS contain flexible functions for workflow management (e.g., Open Text Livelink).

Intranet. The term Intranet denotes an organization-internal ICT platform based on Internet technologies[443]. An Intranet consists of a bundle of applications and data bases. Access to the Intranet is restricted to a limited group of users (also

---

442. Weblogs and Wikis have become popular in the Internet (Wikipedia, Blogosphere). However, many organizations attempt to profit from the benefits of easy content handling also for professional use within the organizational boundaries. Some authors even consider Weblogs and Wikis as (simple) tools for knowledge management (e.g., Efimova 2004, Röll 2006).

443. For an overview of Internet technologies see Röckelein (1999, 22ff). Röckelein uses a model with three layers to describe (1) base technologies, (2) net technologies as well as (3) information services that can be found in public electronic networks such as the Internet. Additionally, he gives a short overview of support technologies and presents numerous examples for the use of Internet technologies for organizations' market communications (Röckelein 1999, 7ff and 109ff respectively). For potentials of an Intranet for businesses see Jaros-Sturhahn/Hießl 1998.

Thiesse/Bach 1999, 105ff). In 1997, one in four German organizations were considered pioneers in the application of Intranets (Jestczemsky 1997, 24). 78% of these pioneers used their Intranet to provide access to data bases, 78% to exchange data and documents, 65% for email, 65% for access to on-line services, 52% for training and education, and 26% for access to financial data, stored e.g., in ERP systems (Jestczemsky 1997, 25).

Groupware. Groupware is a category of software for the support of work groups and teams. Examples for Groupware applications are (Watson 1999, 441f): electronic discussion groups, electronic meeting support, group support systems[444], conferencing software, shared screen systems, group calendars, workflow automation, image management or desktop video conferencing. Groupware is usually classified according to a matrix of group interaction with the two dimensions *time* and *place*: same time vs. different time as well as same place versus different place. Groupware tools can further be classified into (1) communication systems, e.g., email, audio/video systems, chat systems, (2) information sharing systems, e.g., message boards, tele-consultation systems, co-browser, (3) cooperation systems, e.g., co-authoring, shared CAD, whiteboard, word processor, spreadsheet, group decision support systems, (4) co-ordination systems, e.g., group calendar, shared planning, notification systems and (5) social encounter systems, e.g., media spaces, virtual reality (Andriessen 2003, 12). A *Groupware platform* provides general support for collecting, organizing and sharing information within (distributed) collectives of people, such as work groups and project teams over corporate networks as well as the Internet. The best known Groupware platform is Lotus Notes which combines data base, group calendar, email and workflow automation functionality (Watson 1999, 442ff). Other examples are BSCW[445] that is freely available over the Internet and Groove[446], a recent example for a Groupware platform that uses the peer-to-peer metaphor instead of the client-server paradigm.

Data warehousing. A data warehouse is a subject-oriented, integrated, non-volatile, time-variant collection of data in support of management decision processes (Inmon 1992). It is implicitly assumed that a data warehouse is physically separated from operational systems (transaction processing systems, TPS). TPS and also organization-external data bases are the sources from where data are regularly loaded into the data warehouse. Data are organized by how users refer to it. Inconsistencies are removed and data are cleaned (errors, misinterpretations), converted (e.g., measures, currencies) and sometimes summarized and denormalized before they are integrated into the data warehouse (Gray/Watson 1998, 8ff, Muksch/Behme 1998a, 40ff). The data in the data warehouse is usually optimized for the

444. See “Group support systems (GSS).” on page 277.
445. Basic Support for Cooperative Work, offered by the GMD (Gesellschaft für Mathematik und Datenverarbeitung), URL: http://bscw.gmd.de/
446. URL: http://www.groove.net/

use with business intelligence tools (e.g., star and snowflake data model, multidimensional data bases, Gray/Watson 1998, 66ff, Holthuis 1998, 148ff).

Business intelligence. Business intelligence denotes the analytic process which transforms fragmented organizational and competitive data into goal-oriented "knowledge" about competencies, positions, actions and goals of the internal and external actors and processes considered (Grothe/Gentsch 2000, 19). The analytic process requires an integrated data basis that is usually provided by a data warehouse. There are a number of technologies that support this process[447]. Examples are decision support system (DSS) technologies, multidimensional analysis (online analytical processing, OLAP), data mining, text mining and Web mining technologies, the balanced scorecard, business simulation techniques, and also artificial intelligence technologies, such as case based-reasoning or issue management[448],

Group support systems (GSS). GSS are also called group decision support system (GDSS). A GSS is an interactive system that combines communication, computer, and decision technologies to support the formulation and solution of unstructured problems in group meetings[449]. GSS integrate technologies to support the communication in groups, the structuring of processes by which groups interact (e.g., agenda setting, facilitation) and information processing (e.g., aggregating, evaluating or structuring information, Zigurs/Buckland 1998, 319). GSS can be classified according to the level of support in *level 1 GSS* which remove communication barriers, *level 2* GSS which provide decision modeling and group decision techniques and *level 3* GSS which provide expert advice in the selecting and arranging of rules in a meeting and thus lead to machine-induced group communication patterns (DeSanctis/Gallupe 1987, 593ff). One of the best known GSS well received in the literature is GroupSystems (e.g., Valacich et al. 1991, Dennis 1996).

Visualization of structure. Visualization is used in a multitude of tools and systems. Most visualization systems are based on graph theory. In addition to two-dimensional graphs representing elements and relationships, recently a number of tools also provide three-dimensional visualization techniques[450]. Examples are tools for data, function, organization, process or object-oriented modeling or tools that provide mapping techniques which have a long tradition in psychology, sociology and pedagogy, such as mind mapping[451].

---

447. E.g., Gray/Watson 1998, 123ff, Chamoni/Gluchowski 1998, Bissantz et al. 1998, Watson 1999, 469ff, Grothe/Gentsch 2000, 21.
448. See "AI technologies." on page 279.
449. DeSanctis/Gallupe 1987, 589, Turban et al. 1996, 501, see also Zigurs/Buckland 1998, 320 for an overview of classifications of GSS technologies.
450. So-called hyperbolic browsers, see also section 7.4.5 - "Collaboration services" on page 327.
451. See e.g., Mandl/Fischer 2000 for an overview of mapping techniques which can be applied in knowledge management.

Search engines. A search engine is a program that can be used to find Web sites, documents or images, either in an organization's Intranet or in the WWW. Search engines apply programs that permanently trace the Web or an Intranet for new Web pages, so-called spiders or robots (Horn 1999, 57, Brenner et al. 1998, 197ff). A new found Web page is scanned for possible keywords which then are stored together with the URL of the Web page in the search engine's data base. At the time when a user submits a search term to the search engine, only this data base is searched and intelligent algorithms are applied in order to retrieve those Web pages that fit most to what the user has searched for. One of the best known search engines that is used in a number of KMS is Verity's K2 Enterprise or Developer search engine[452]. So-called meta- or multi-search engines (Horn 1999, 59) forward search strings including boolean operators to various search services, collect and filter the results (e.g., for redundancies) and present them accordingly. One of the best known meta-search engines on the Internet is Meta-Crawler[453]. Both, search engines and meta-search engines can be further distinguished with respect to the search domain which they support, such as organization-internal and/or organization-external systems.

Computer based training (CBT) tools and learning environments.
Learning environments are application systems that offer specified learning content to the learner in an interactive way and thus support the teaching and/or learning process (Behrendt 1998, 220, Schäfer 2000, 36). CBT, also called *computer-assisted* or *aided instruction* (CAI) or *computer supported learning* (CSL)[454], has its historical roots in *programmed instruction* or *learning* in the late 1950s and 1960s which was based on the concept of *operant conditioning* developed by Skinner (Hilgard/Bower 1975, 610ff, Möhrle 1996, 76ff). Both, psychological and pedagogical as well as technological advancements have led to a wide variety of CBT systems and learning environments which reflect how diverse learning can be[455]. Examples are[456]: drill & practice systems, (intelligent) tutoring systems, active assistance systems, microworlds, simulation systems, experimental game systems, hypertext-/hypermedia learning systems as well as more recent developments in the field of computer-supported learning, such as Web based training (WBT), multimedia learning environments, tele-teaching, distance learning, tele-tutoring and computer supported collaborative learning. Recently, these diverse CBT concepts

452. See URL: http://www.verity.com/; see also the support Web site for this book http://iwi.uibk.ac.at/maier/kms/.
453. URL: http://www.metacrawler.com.
454. There are many more terms in use that denote the application of software for teaching and/or learning purposes (e.g., Bodendorf 1990, 37f) which reflects the vivid interest in this field, especially since the 80s and the wide-spread use of the PC.
455. For examples see Schanda 1995, 21ff, Ballin/Brater 1996, 41ff, Möhrle 1996, 24f, Schulmeister 1997.
456. See also Möhrle 1996, 32ff, Mertens et al. 1997, 46, Behrendt 1998, Kerres 1998, Schreiber 1998, 11ff, 16f, Lehner/Klosa 2000, Schäfer 2000, 38ff, Lehner 2001, Nikolaus 2002, 22ff.

have found their way into integrated learning management systems or e-learning suites which overlap with KMS[457].

Communication systems. Communication systems are electronic systems that support both asynchronous and synchronous communication between individuals (point-to-point communication systems) and collectives (multi-point communication systems). Examples for synchronous communication systems are tele-conferencing systems such as text conferencing (chat), instant messaging, audio and video conferencing systems. Examples for asynchronous communication systems are email, listserver or newsgroups[458].

AI technologies. There are a large number of specific technologies that is discussed as supporting knowledge management. Most of these technologies have their roots in the field of artificial intelligence. Results from AI research play a crucial role in the development of KMS and provide intelligent functions for KM. Examples for AI-based tools for KM are[459]:

- *experience and know-how data base systems* are ordered collections of application solutions, i.e., specialized data base systems that store e.g., experiences, lessons learned, best practices as well as technical solutions (Mertens et al. 1997, 227f, Roithmayr/Fink 1997, 503). Experience data bases technologically typically rely on conventional information retrieval and document management technology, augmented with business process models and ontologies about the application domain as well as additional meta-data categories for describing knowledge documents (Kühn/Abecker 1997, 932, Staab et al. 2001). The term experience data base aims more at management, organizational and technical experiences (e.g., customer relations, business processes, projects) whereas the term know-how data base aims more at technical problems and solutions (Wargitsch 1998, 25f);
- *case-based reasoning (CBR) systems* provide an approach to solve problems with the help of known solutions for similar problems that has its roots in AI research. CBR comprises the four steps (1) retrieve cases from the system's case base which are similar to the problem presented by the user, (2) reuse solved cases, (3) revise the selected case and confirm the solution and (4) retain the learned case if it is an interesting extension of the case base (Aamodt/Plaza 1994[460]);

457. See section 4.3.1 - "Overview and related concepts" on page 82.
458. See also section 7.4.5 - "Collaboration services" on page 327.
459. See also Kühn/Abecker 1997, 931ff, Mertens et al. 1997, Probst et al. 1998, Wargitsch 1998, 23ff, Krallmann et al. 2000, 234ff, Lehner 2000, 330ff, Mertens/Griese 2002, 49ff.
460. For an extensive analysis and discussion of the potentials of CBR see also Althoff/Aamodt 1996, Mertens et al. 1997, 74f, the special issues on case-based reasoning of the journal WIRTSCHAFTSINFORMATIK, Ehrenberg 1996 or the journal KI, Bartsch-Spörl/Wess 1996; examples of CBR tools are listed on the support Web site for this book http://iwi.uibk.ac.at/maier/kms/; see also the overview of CBR tools and applications, URL: http://www.cbr-web.org/.

- *recommender systems* extend systems that support information retrieval and give recommendations based on techniques such as test of context correspondence, frequency analysis and agent technologies (e.g., Wargitsch 1998, 29). Some authors also use the term *collaborative filtering* (Goldberg et al. 1992) to denote the social process of recommending. The systems collect and aggregate recommendations of a multitude of people and make good matches between the recommenders and those who seek recommendations (Resnick/Varian 1997, 56). In order to accomplish this, recommender systems have to model the users' characteristics, interests and/or behavior: *user modeling* (Bodendorf 1992, Mertens/Höhl 1999), also called *profiling* (Brenner et al. 1998, 132ff, Applehans et al. 1999, 37ff) or *personalization* (Zarnekow 1999, 132f). Profiles are a requirement for the application of many intelligent technologies, especially intelligent software agents (see next paragraph). Systems using *content-based filtering* recommend items similar to those a given user has liked in the past (Balabanovic/Shoham 1997, 66). Recently, AI techniques as part of recommender systems have been applied widely in commercial Web sites, e.g., to recommend music, videos or books (e.g., URL: http://www.amazon.com/[461]).
- *intelligent software agents* are autonomous units of software that execute actions for a user (Mertens et al. 1997, 6). Intelligent software agents use their intelligence to perform *parts of its tasks* autonomously and to interact with its environment in a useful manner (Brenner et al. 1998, 21). Software agents thus differ from more traditional software programs with respect to their autonomy, ability to communicate and cooperate, mobility, reactive and proactive behavior, reasoning, adaptive behavior and last but not least some agents even might show human characteristics (Zarnekow 1999, 16ff). The roots of the agent technology can be traced back to approaches of *distributed artificial intelligence* where agents deconstruct tasks into sub-tasks, distribute the sub-tasks and combine their results (Mertens et al. 1997, 7) and to developments in the area of networks and communication systems which form the underlying technological basis (Brenner et al. 1998, 41f). Intelligent or semi-intelligent agents can be classified according to their main area of application as information agents, cooperation agents and transaction agents (Brenner et al. 1998, 19) and are applied in a multitude of settings. Prominent examples for agents can be found in electronic market processes. Agents provide value-added services for the identification phase, the information phase, the negotiation and buying phase (in a narrow sense) as well as the application and service phase of a buying process (Zarnekow 1999, 118ff). In knowledge management, agents can be used e.g., to scan emails, newsgroups, chats etc., to group and automatically update user-specific messages and information items in the Internet (newswatchers), to analyze

---

461. For a more detailed discussion and examples of recommender systems see Konstan et al. 1997 (GroupLens; for netnews articles), Kautz et al. 1997 (ReferralWeb; for people), Terveen et al. 1997 (PHOAKS; for URLs) and Rucker/Polanco 1997 (Imana's CommonQuest; for URLs).

and classify documents, to search, integrate, evaluate and visualize information from a multitude of sources, to intelligently handle information subscriptions, to identify and network experts, to visualize knowledge networks and to recommend participants, experts, communities and documents[462].

- *issue-based information systems* are systems to visualize argumentation that build structured networks of arguments consisting of e.g., questions, opinions, pro and counter-arguments or examples recorded in group decision processes (Buckingham Shum 1998, 903ff, Wargitsch 1998, 29). One of the best known examples is the system gIBIS which is marketed as CM/1 or QuestMap respectively (Conklin/Begeman 1988, Stein/Zwass 1995, 93, Buckingham Shum 1988, 906ff).

7.2 Contents

The content of an organizational memory—the organization's knowledge—can be located[463]:

- in *peoples' minds*,
- in *artifacts*, such as the physical organization, e.g., the architecture, the use of office space; printed media, audiovisual media and multimedia instruments etc.,
- in *ICT systems*, particularly in KMS, e.g., routines, procedures, models, (hypertext) documents, multimedia files, user profiles, learning (CBT) modules, knowledge bases or links to experts.

These three locations, or media, are related to each other and complexly interwoven into knowledge networks. Networks of knowledge consist of a number of people with their external memories, e.g., documents, office space and ICT systems. These networks of knowledge have been termed organizational competencies which in turn create competitive advantages[464]. Consequently, KM has to handle and improve these complex relationships and networks rather than individual knowledge elements or just one location, e.g., a knowledge base. The transactive memory system concept (Wegner 1986) has been suggested to analyze these complex relationships and provides a great metaphor for the implementation of KMS and especially for structuring the contents.

Due to the complexity of this topic and the focus of this book the following discussion of contents will concentrate on KMS[465]. Generally, both, normative sug-

---

462. For examples of actual implementations of some of these technologies see Brenner et al. 1998, 189ff, Zarnekow 1999, 163ff and the list of KMS provided on the support Web site for this book http://iwi.uibk.ac.at/maier/kms/.
463. See also Watson 1999, 15 who concentrates on people and electronic organizational memories and Amelingmeyer 2000, 51ff who distinguishes between persons, material media and collective media as locations for knowledge. The idea of a collective or organizational memory is discussed in section 4.1.1 - "From organizational learning to knowledge management" on page 22; different types of knowledge including collective knowledge are investigated in section 4.2.2 - "Types and classes of knowledge" on page 66.
464. See section 5.1 - "Strategy and knowledge management" on page 93.

gestions for KMS and actual implementations of KMS, vary considerably in terms of the content to be managed. According to the interviews with knowledge managers, many companies seem to be driven by a pragmatic approach which puts those parts of the organizational knowledge at the center of consideration the management of which promises the most direct positive effects. Typically, the organization's knowledge structure is determined in a workshop and reflects sources that already exist in the organization, at best in electronic form, but are handled by a number of incompatible ICT systems. Examples are customer-related data, patents, skills data bases (yellow pages), lessons learned, best practices, descriptions of products, business processes, the structural organization or projects, external online data bases, presentations, reports and market studies. In many cases, explicit knowledge is predominant. It is also a lot harder to describe implicit knowledge that is an equally important part of knowledge to be handled in organizations.

Section 7.2.1 discusses examples for *types of contents* that can be found in KMS. Section 7.2.2 defines the concept of a knowledge element and discusses some aspects of maturity of knowledge. Section 7.2.3 investigates what *media formats* are supposedly used to encode knowledge elements and how to determine the size of organizational knowledge bases. Finally, section 7.2.4 discusses the two predominant ways to *organize knowledge elements*, the hierarchical and the network structure.

7.2.1 Types of contents

A classification of types of contents of KMS can be built on the abundance of classifications and distinctions of types of knowledge as presented in section 4.2.2 - "Types and classes of knowledge" on page 66. Some pragmatic distinctions which can be studied rather easily are:

- *organization-internal*, that is knowledge created inside the organization, e.g., internal analysis, versus *organization-external knowledge*, e.g., market reports,
- *formal knowledge*, that is knowledge already approved by some institution and officially released, e.g., descriptions of organization and processes, versus *informal knowledge*, e.g., ideas, questions and answers,
- *secured knowledge*, that is knowledge protected by intellectual property right or some other form of legal contracts, e.g., patents, versus *securable knowledge*, e.g., a part of proposals or best practices, versus *knowledge not securable*, e.g., external patents, common industry knowledge,
- *historic knowledge*, that is knowledge that relates to past events, experiences or has been used in a certain application context, e.g., lessons learned, versus *knowledge relating to the future*, that have not been used in the past, but have a prescriptive or normative character, e.g., proposals, ideas.

---

465. Research about knowledge processing and representation in people's heads has a long tradition in the field of cognitive psychology (see also section 4.1.1 - "From organizational learning to knowledge management" on page 22). Architecture has been briefly touched in section 6.5 - "Other interventions" on page 230.

- *classification according to the topic*, e.g., knowledge about participants, customers, business partners, stakeholders, competitors, products, methods, instruments or procedures[466].

In order to get a more detailed picture, a list of sixteen items will be used that represent typical contents of KMS in the empirical study. The list was pragmatically developed on the basis of the literature and several interviews with knowledge managers. There are two different theoretical streams that were used for the classification of the type of contents of KMS. These are the distinctions between:

Integrative and interactive KMS[467]. It is supposed that the predominant knowledge managed in *integrative KMS* currently will be *method, product and process knowledge* whereas in *interactive KMS* the main knowledge used will be *person-oriented knowledge*.

Novices and experts. The classification distinguishes between knowledge adequately presented for *novices*, i.e. facts and rules, and knowledge better suited for the perception by *experts*, i.e. case-oriented knowledge, or at least competent[468]. This is a differentiation well-suited to detail both, method, product and process knowledge as well as person-oriented knowledge.

Table B-13 shows some examples for each type of knowledge which will be described in the following[469].

Knowledge about organization and processes. Descriptions of the organization (structure and processes) are typically managed by the IT/organization, HRM departments or by process owners and managers. Examples are organizational charts, event-driven process chains to describe business processes, descriptions of organizational positions, projects, roles or personnel handbooks.

Product knowledge. This type of knowledge represents descriptions related to the organizations' products and/or services, such as marketing presentations, technical papers, CAD models or white papers.

---

466. For an extensive list of dimensions of types of knowledge see section 5.2.2 - "Strategic options" on page 120.
467. See section 7.6.1 - "Knowledge Tools" on page 361.
468. See section 6.1.2 - "Knowledge management roles" on page 162 for a discussion of novices versus experts.
469. Several types of knowledge described in the following are specifically targeted by a corresponding KM instrument, e.g., lessons learned, good or best practices etc. which have been described in section 6.2 - "Instruments" on page 195. The number of types of knowledge does not amount to sixteen as some types have been split in the table, i.e. studies and business partners, as well as combined in the following descriptions, but treated separately in the empirical study, e.g., patents and studies. The latter thus amount to fifteen types of knowledge to which private contents are added so that there are sixteen types of knowledge that have been tested in the empirical study in 14.2.1 - "Types of contents" on page 532.

Internal/external patents. Patents are legally secured innovations. There will be a distinction between patents held by the organization and organization-external patents. External patents can be found e.g., in so-called patent data bases such as the World Patent Index (WPI, operated by Derwent, Mertens/Griese 2002, 22).

TABLE B-13. Classification of knowledge with respect to type and target group

	method, product and process knowledge	**person-oriented knowledge**
facts and rules (novice)	• knowledge about organization and processes • internal and external patents • product knowledge • fact knowledge in internal/ external studies and analyses	• employee yellow pages • fact knowledge about business partners
cases (expert)	• lessons learned • best practices • ideas, proposals • cases in internal/external studies and analyses	• cases about business partners • directory of communities • employee communication • questions, answers (frequently asked questions, FAQ)

Internal/external studies/analyses. Reports document the results of an organization-internal study or analysis related to a specific topic or a study or analysis performed by an organization-external institution, e.g., universities, research institutions, professional services companies or benchmarking groups.

Lessons learned. Lessons learned are the systematically documented essence of experiences made by members of the organization in e.g., projects or learning experiments. They thus are authored by a collective of project members that commit to the critical experiences made in the project and documented for future reuse in the same or in other projects.

Best practices. This term in a wide meaning denotes knowledge in a process-oriented form that describes task or workflows that have proven to be valuable or effective within one organization or organizational unit and may have applicability to other organizations (also O'Dell/Grayson 1998, 167). Regularly, best practice management distinguishes various categories of quality that relate to the scope in which the corresponding practice is considered "best", e.g., team, subsidiary, company; group[470] or industry best practice (O'Dell/Grayson 1998, 167).

Ideas, proposals. These can be informal or formal documents submitted to an established proposal system. So-called *microarticles* are a structured approach to

470. In the sense of a group of companies belonging to the same concern, e.g., the BMW Group.

organize individual learning experiences and help knowledge workers to externalize and share their knowledge (Willke 1998, 107ff).

Questions, answers (FAQ). Frequently asked questions (FAQ) are a popular instrument to store questions that might be of interest to many participants together with answers, mostly given by experts (e.g., Mertens/Griese 2002, 52). Examples are the manyfold public FAQ lists that can be found in newsgroups or the WWW.

Employee yellow pages. Expert yellow pages and skills directories support the transparency of expertise in an organization. Employees can provide their skill or competence profile which can be accessed by all employees who look for an expert on a certain topic or for an expert who can provide a solution to a given problem.

Knowledge about business partners. This topic-specific type of knowledge has been gained from interactions with customers and suppliers, e.g., through personal or computer-supported interaction between business partners and members of the organization, customer relationship management, supply chain management programs and surveys.

Directory of communities. In analogy to skills directories, this is a list of communities that are established within or accessible through the organization and a short description of themes, members and contact data. The directory might also offer some examples for discussions that are mediated or for documents that are shared with the help of community home spaces.

Internal communication. This term denotes the organization-internal equivalent to public relations and describes the part of corporate communication that is targeted to the organization's employees: official organization-wide communication, e.g., business TV, corporate newsletters, corporate electronic magazines, announcements etc.[471].

External on-line journals. The electronic equivalent to paper-based journals can be directly accessed through the Web[472]. Due to the fact that on-line journals can hold both types of knowledge as well as fact knowledge and cases, they cannot be classified according to the dimensions in Table B-13 on page 284.

Organizations with a systematic KM initiative supposedly handle different types of knowledge when compared to organizations without such an initiative. The list of items describing the contents of KMS contains several items which require special attention in order to be systematically handled in the organizations' electronic knowledge bases. These are best practices, lessons learned and employee yellow

471. See Will/Porak 2000, 195f for an extensive model of corporate communication that covers internal and external communication.
472. For example the Knowledge Management Magazine, URL: http://www.kmmagazine.com/.

pages. Moreover, at many KM conferences organizations that handle knowledge that is legally secured (patents) were on the forefront of applying KM (e.g., chemical or pharmaceutical organizations). Again, this points to the direction that organizations with systematic KM differ from other organizations with respect to contents handled in their KMS. The following hypothesis will be tested:

Hypothesis 13: Organizations with systematic knowledge management target different contents than organizations without such an initiative

In addition to the 15 items describing the contents of KMS, *private contents* were included as it is hypothesized that this in turn has significant effects on the way an organization handles knowledge. By allowing employees to publish private contents or to present themselves, organizations can show that they respect the individuals' off-the-job interests and networking needs. If organizations take these needs and interests seriously, it might in turn have a positive influence on the building of trust and as a consequence the willingness to share knowledge of their employees.

Hypothesis 14: If an organization allows private contents as part of their knowledge management systems, willingness to share knowledge is higher

7.2.2 Maturity of knowledge elements

The term content and its treatment with the help of ICT takes an objectified perspective on knowledge[473]. A knowledge unit or *knowledge element*, sometimes also called knowledge chunk, denotes the smallest unit of explicit, documented knowledge. It has been termed "a formally defined, atomic packet of knowledge content that can be labeled, indexed, stored, retrieved, and manipulated. The format, size and content of knowledge units may vary, depending on the type of explicit knowledge being stored and the context of its use" (Zack 1999a, 48). Examples for knowledge elements are (Zack 1999a, 49):

- concepts, categories and definitions (declarative knowledge),
- processes, actions and sequences of events (procedural knowledge),
- rationale for actions or conclusions (causal knowledge),
- circumstances and intentions of knowledge development and application (specific contextual knowledge).

However, these are still conceptual categories. From an ICT perspective, examples for knowledge elements are:

- a document, email message, instant message, video file, audio file, slide show or picture displaying an idea, proposal, recommendation, an expert's opinion, a description of or solution to a specified problem[474],
- a personal note with a write-up of a personal experience,

473. See section 4.2 - "Knowledge" on page 60, particularly the discussion related to the description of Figure B-8, "The term knowledge and its application in KM," on page 78.

- a contribution to a forum, newsgroup, Wiki, Weblog or other form of CMS,
- an entry in a list of frequently asked questions (FAQs) and the answer to the question,
- an element in an experience data base,
- a document with e.g., a product presentation, lesson learned, good or best practice, story, study, write-up of an experiment, whitepaper, patent or report, e.g., about the results of a project milestone,
- a prototype,
- a model of e.g., a (business or knowledge) process, class, data, knowledge structure or other enterprise model,
- a learning object in a learning repository, e.g., definition, explanation, formula, example, case, demonstration, exercise, exam question, test or master solution,
- a skill description in a skill data base,
- an entry in a yellow page system or expertise locator describing available expertise on a specified topic,
- knowledge elements that connect some of the above elements to persons, groups, teams or organizational units, e.g., the description of skills of a particular employee or organizational unit,
- an evaluation of or a comment to one of these knowledge elements etc.

The types of data underlying these knowledge elements have been extended from structured data as can be found in data base systems to (semi-)structured data typically found in e.g., DMS, file servers, CMS or email servers. As compared to structured data, semi-structured data has not been managed equally well in most organizations. A large number of terms have been coined for semi-structured data, e.g., content, (digital) asset or, most importantly for the handling of knowledge elements, the term document.

A document is a legally sanctioned record or a transitory record of a business transaction, decision or some form of externalization of knowledge that can be viewed as a single organized unit both from a business or knowledge perspective and from a technical perspective. It is composed of a grouping of formatted information objects which cannot be separated without substantial loss of meaning, possibly together with meta-data[475].

The term record denotes that the document's context relates to some kind of business transaction or decision or, in the case of knowledge elements, some form of externalization of knowledge, which the document represents. Examples for legally sanctioned records are purchase orders or patents. Examples for transitory

474. The stress is here on the representation of a solution to a specified problem. This is not necessarily a document, a video file or an audio file etc., but can also be a selected portion, e.g., a document fragment, a video sequence or an audio theme.
475. See also Kampffmeyer/Merkel 1997, 1999, Karakas 2003, Götzer et al. 2004, Maier et al. 2005, 247ff, Maier/Trögl 2006.

records are meeting notes or ad-hoc solutions to problems. There are legal requirements and retention plans regulating the handling of many types of documents in organizations, e.g., access restrictions or time period required for archival. The term transitory reflects the fact that not all documents are archived, but some are developed step-by-step with increasing levels of maturity which calls for versioning. Documents are collections of information objects bound by the document's purpose. These information objects are often formatted, so that in some cases, e.g., certain contracts or annotated maps, the original form of the entire document has to be conserved. Documents can be regarded as containers of content which cannot be split without loosing their original meaning and, in the case of knowledge elements, without loosing context and thus hindering reconstruction of knowledge. Annotations with meta-data ease transfer, distribution, retrieval and understanding of documents[476]. Documents are accessed as a whole because they group related information with respect to the expected or most common user needs.

Documents can be *elementary*, e.g., a text file or a fax message, *compound*, e.g., a text file with embedded graphs, tables or pictures or *container*, e.g., a collection of elementary or complex documents organized around a workflow in a folder or zip file (Kampffmeyer/Merkel 1997, 12). Documents have business value and thus can be considered as (digital) assets. Document types can be distinguished using a number of characteristics, for example:

- physical characteristics, primarily with respect to non-electronic documents,
- formal characteristics, e.g., file types and formats,
- structure, e.g., functional grouping of objects, sequence,
- type of content, e.g., type of knowledge element,
- layout, e.g., arrangement, design,
- coding, coded or non-coded information,
- time characteristics, e.g., date of creation, last modification, last access, version,
- control and security characteristics, e.g., encryption, confidentiality, privileges to search, access, print, change, create, delete or administer documents,
- legal characteristics, e.g., requirements for retention, modifiability, digital rights management.

Taking into account the definition of document, Box B-8 defines the term knowledge element. The considerable variety of (1) types of knowledge elements, of (2) organizational units responsible for a systematic management of the processes in which these knowledge elements are involved as well as of (3) systems supporting these knowledge elements leads to an often fragmented landscape of numerous media and locations to preserve as well as channels to transfer knowledge of varying degrees of maturity which employees, teams, work groups and communities can select from in order to retain or transfer knowledge elements for further development and application by other employees, teams, work groups or

476. See section 7.7.2 - "Meta-data management" on page 379.

communities. The choice is often difficult, leading to inadequate supply of information and knowledge in organizations and thus can be improved.

A knowledge element is the smallest unit of atomic, explicit, formally defined knowledge content, a record of some form of externalization viewed as a single organized unit both from a conceptual and from a technical perspective. It is composed of a grouping of formatted information objects which cannot be separated without substantial loss of meaning together with meta-data describing the element.

BOX B-8. **Definition of knowledge element**

Examples for types of knowledge elements have been given in section 7.2.1 - "Types of contents" on page 282. Organizational units, such as innovation management, project management, quality management or units dealing with e-learning, all intend concurrently to improve construction, preservation, integration, transfer and (re-) use of knowledge and competencies. Additionally, programs of personnel development as part of HRM support training into the job, on the job, near the job, off the job and out of the job (Scholz 2000). But despite increased interest in bringing them together, particularly as part of KM initiatives, there are still huge conceptual differences. Whereas e-learning and personnel development have their foundations in (learning) psychology, (media) didactics and (learning) pedagogy and emphasize the importance of structural guidance by preparing learning material or personal guidance, there are also more document-oriented units, such as project and quality management that rather envision an organizational knowledge base into which the individual's knowledge is supposed to be made explicit and which is the basis for more or less unguided knowledge transfer.

From an ICT perspective, numerous systems aim at improving knowledge and learning processes as well as organizational competency development which are typically designed and managed according to the specific needs of the respective organizational units. Employees thus use a fragmented systems landscape in which each system supports a certain part of knowledge and learning processes. There are substantial conceptual challenges of designing learning and knowledge processes that bring together the separated organizational support infrastructures fostered by the different organizational units. Therefore both, organizational units and corresponding application systems typically target knowledge of different degrees of maturity.

Pruning the tree of types of knowledge elements and guiding employees on how to use the channels of knowledge transfer is thus a pivotal task in any KM initiative. In the following, the knowledge maturing process is described in order to provide a framework for the design of the required integrating types of knowledge elements, knowledge processes and channels in KM.

In a first step of structuring this process, Figure B-48 shows the five phases that have been identified after analyzing some practical cases[477]. The phases are described in the following.

- *expressing ideas*: New ideas are developed by individuals in highly informal discussions. The knowledge is subjective and deeply embedded in the context of the originator. The vocabulary used for communication is vague and often restricted to the person expressing the idea.
- *distributing in communities*: This phase accomplishes an important maturing step, i.e. the development of common terminology shared among community members, e.g., in discussion forum entries or Blog postings.
- *formalizing*: Artefacts created in the preceding two phases are inherently unstructured and still highly subjective and embedded in the context of the community. In this phase, purpose-driven structured documents are created, e.g., project reports or design documents in which knowledge is desubjectified and the context is made explicit.
- *ad-hoc learning*: Documents produced in the preceding phase are not well suited as learning materials because no didactical considerations were taken into account. Now the topic is refined to improve comprehensibilty in order to ease its consumption or re-use. The material is ideally prepared in a pedagogically sound way, enabling broader dissemination.
- *formal training*: The ultimate maturity phase puts together individual learning objects to cover a broader subject area. As a consequence, this subject area becomes teachable to novices. Tests and certificates confirm that participants of formal training have achieved a certain degree of proficiency.

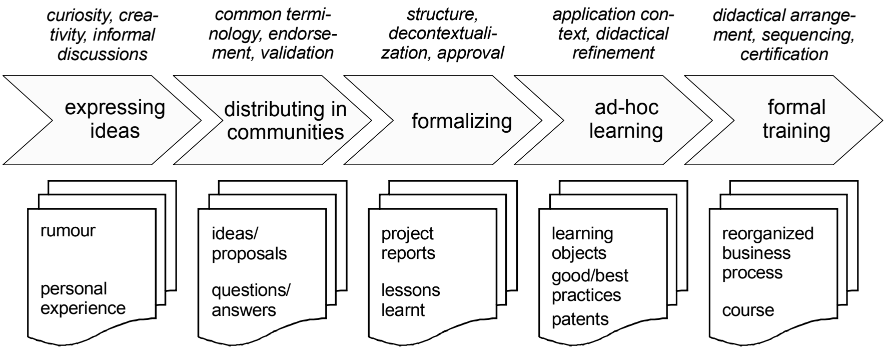

FIGURE B-48. Knowledge maturing process[478]

477. See Schmidt 2005, Maier/Schmidt 2007 who considered project experiences as reported in Bayer et al. 2005, Schmidt/Braun 2006 as well as metaphors of organizational knowledge and learning discussed in chapter 6 - "Organization" on page 153 and also the empirical results on types of contents presented in section 14.2.1 - "Types of contents" on page 532.

Knowledge thus can be classified according to its level of maturity. The class then suggests the appropriate form of learning and technical support systems. The following criteria have been identified as useful to define classes of knowledge:

Validity. Certainly, the most obvious categorization refers to a validation process of knowledge and could distinguish in a first step between unproven and proven[479] knowledge. In a more refined version that considers the specifics of organizational knowledge, validation could take into account the number of successful uses of knowledge, systematic tests or, finally, (mathematical) proves for its working.

Hardness. In analogy to mineralogy, this criterion describes the (alleged) validity and reliability of information or knowledge. According to Watson (1999), a possible scale runs from unidentified sources of rumors up to stock exchange data (see Table B-14).

TABLE B-14. Scale for information hardness[a]

degree	description	degree	description
1	unidentified source; rumors, gossip and hearsay	6	budgets, formal plans
2	identified non-expert source; opinions, feelings, ideas	7	news reports, non-financial data, industry statistics, survey data
3	identified expert source; predictions, speculations, forecasts, estimates	8	unaudited financial statements, government statistics
4	unsworn testimony; explanations, justifications, assessments, interpretations	9	audited financial statements, government statistics
5	sworn testimony; explanations, justifications, assessments, interpretations	10	stock exchange and commodity market data

a. Source: Watson 1999.

478. After: Maier/Schmidt 2007. When comparing this basic model with the model of organizational information processing (see Figure B-22 on page 154), all processes in the basic model are also part of the model of information processing. The emergence of ideas corresponds to the process of individual learning, distribution in communities corresponds to sharing, formalization is reflected in institutionalization, ad-hoc training in feedback and formal training in the refining and repackaging processes. The basic model sets the focus on a pragmatic chain of knowledge development tasks that can be designed so that formal, mature knowledge products are the outcome of the respective knowledge maturing process.

479. In a critical-rationalist perspective, "proven" could be replaced by repeatedly not falsified. It is noted that validation or "truth" of knowledge is a category that gives rise to age-old philosophical debates which this book will refrain from; for a small account see section 4.2 - "Knowledge" on page 60.

Context. With deepened understanding, connections to other topics become visible. This must not be confused with inherent contextualization of knowledge which decreases in the knowledge maturing process and refers to the degree of implicit linkage to the creation context, so that it cannot be used outside the original context. Inherent contextualization and inter-connectedness are inverse properties.

Commitment/legitimation. Knowledge can be structured according to the amount of support it gets. Support can be in the form of commitment by members of groups, teams, communities or other organizational units. Another form of support can be authorization to use knowledge by supervisors, executives or committees as well as legalization and standardization, forms of legitimation (Figure B-49).

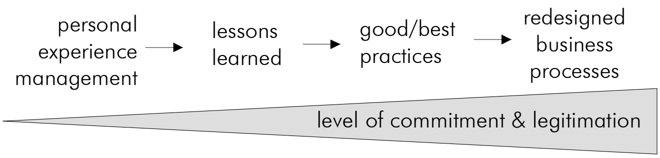

FIGURE B-49. Portion of the knowledge life cycle

The knowledge life cycle starts with individual experiences which have the least level of organizational commitment. Individual experiences are discussed, filtered and further explored in a team. If the team commits to certain experiences, they are called lessons learned. This process can be aided by a lessons learned coach that helps the team to structure the process of group reflection on team experiences. Further commitment and legitimation is needed in order to turn lessons learned into good practices. Practices can be seen as guidelines how to act in certain situations. Sharing good practices throughout the organization and agreeing that this is the best way to deal with a specific situation turns them into (organizational or local) best practices. Knowledge process reengineering is finally one method for redesigning business processes taking good and best practices into account. Knowledge bound to an individual is disseminated in the form of knowledge products that ultimately reside in social systems, changed business practices and processes.

Form of learning. As knowledge maturing is basically interconnection of individual learning processes where knowledge is taught and learnt, an important criterion is its teachability. Whereas immature knowledge is hard to teach (even to experts), formal training allows by definition for wide-range dissemination.

Table B-15 gives an impression of what a checklist for the classification of knowledge elements according to the criteria for maturity of knowledge discussed above could look like. This exemplary list differentiates between the four maturity levels initial, advanced, consolidated and mature. The last three rows give examples for types of knowledge and learning objects as well as channels that could be institutionalized to capture knowledge of varying degrees of maturity. The checklist should help organizations to design supporting infrastructures for maturing knowledge. These infrastructures are thought of as both, organizational and technical infrastructures. These help to (semi-)automatically identify knowledge that is ready to be brought to the next level of maturity. The knowledge is visualized together with its context in the same maturity level as well as the context of knowledge elements in the next maturity level. Then, the infrastructure could recommend specific actions on the knowledge elements, e.g., selection of certain parts, summaries, tagging, merging or other foms of enrichment and integration.

TABLE B-15. Exemplary categories for maturity of knowledge

criterion	initial	advanced	consolidated	mature
validation	unproven	successfully used	systematically tested	proven
hardness	proposed	supported	approved	audited
context	isolated	filed	annotated/ tagged	linked/ networked
commitment	opinions in community	convergence of discussions	consensus	commitment
legitimation of knowledge	ad-hoc order	guideline	standard operating procedure	compliance to standard
legitimation of learning content	case write-up	peer-reviewed article	textbook by field expert	standard textbook
legitimation of personal advice	peer advice	community advice	company expert advice	field expert advice
teachability	no special attention	explication of learning goals	sequencing	personalization
knowledge type	idea	lesson learned	good practice	patent/process
learning resources	learning material	learning object	course	certified/ personalized course
channel	individual communication	emerging social network	community of practice/interest	centre of competence

Table B-16 gives an overview of the phases of the knowledge maturing process with an examplary list of characteristic types of knowledge and their values according to the criteria discussed in this section. The degree of hardness of types of knowledge is not a direct translation of the scale of information hardness, but attempts to match it as closely as possible. Information hardness only considers individuals and institutions as sources of information, but does not consider teams and communities. In the latter cases, the degree of hardness is thought of as being in between individuals (information hardness 1-5) and institutions (information hardness 6-10). In the case of reorganized business processes, those compliant to laws, regulations and standards are considered of higher hardness. The same applies to courses when they are certified by some external authority.

TABLE B-16. Types of knowledge in phases of knowledge maturing process

phase	knowledge type	hard-ness	medium/context	commitment/ legitimation	form of learning/ technology
expressing ideas	rumors	1	human, highly contextualized	none	informal, direct communication by phone, instant messaging, email
	personal experiences	2	human, personal notes, highly contextualized	commitment by individuals, confirmation by colleagues	direct/computer-mediated communication, exchange of personal artefacts, collaboration systems, Weblogs
distributing in communities	ideas and proposals	2	forum entry, suggestion form, explicit use context	commitment by individuals, confirmation by colleagues	committee selection, validation, organizational proposal system, forum, community workspace
	questions/ answers	3	FAQ, forum entry, explicit problem context	legitimation by experts	self-managed, on-demand search, FAQ data base, forum, Wikis
formalizing	project results	3	project/milestone report, explicit project context	legitimation by project manager	on-demand search, project & document management system
	lessons learned (LL)	4	LL document, explicit project context	legitimation by project team	case-based, self-managed learning, LL data base, Wikis, Weblogs
ad-hoc learning	learning objects	5	well-defined digital resource, formal meta-data	legitimation by experts	self-managed ad-hoc learning, composition from learning object repository
	good/best practices (BP)	5	BP document, process description, explicit creation context	commitment of team, unit, company, group, industry	case-based, self-managed ad-hoc training, continuous process improvement, BP data base
	patents	9	patent application, explicit potential use context	legitimation by patent office	specialized information seeking, patent data bases
formal training	reorganized business process (compliant)	6 (7)	process models and descriptions	legitimation by process owner	standard training of standard operating procedures, courses, process warehouse
	courses (certified)	6 (7)	composed learning objects, curriculum, certificates	legitimation by course owner	standardized training, WBT authoring, learning content management system

Figure B-50 reviews the diagram classifying KM instruments presented in Figure B-24 on page 199. The arrows connecting KM instruments represent some examples for maturity paths between KM instruments that could be systematically designed and encouraged in organizations. The Latin numbers (I-III) show the two major directions in which maturity paths can be organized in organizations:

- from personal-product knowledge via personal-process to organizational process knowledge and
- from personal-product knowledge via organizational-product to organizational process knowledge.

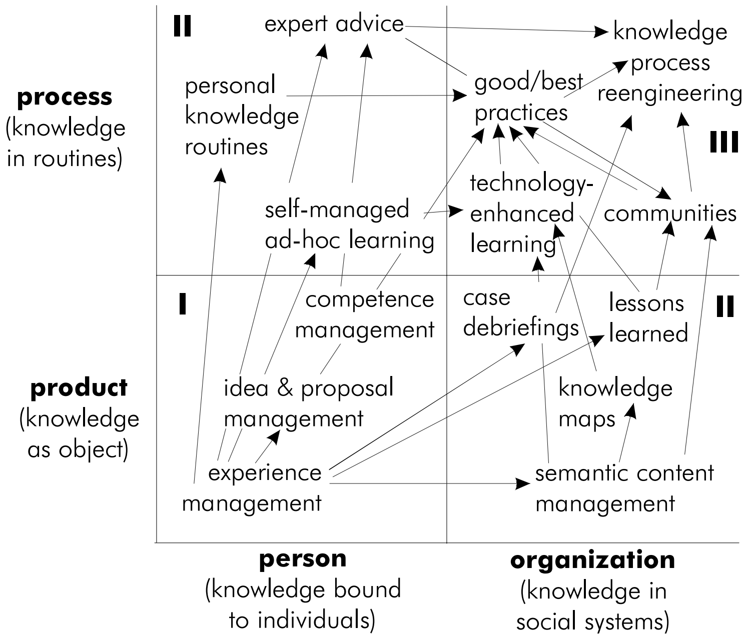

FIGURE B-50. Exemplary maturity paths between KM instruments[480]

However, the maturity path between *idea and proposal management* and *good/best practices* shows that there are also paths that directly relate personal-product with organizational process knowledge. Other paths are thinkable, but have been omitted for reasons of readability. The model can be used by organizations (1) for checking what processes, procedures, roles and system services they have established in each of the categories, (2) for connecting these with the help of explicitly designed transitions along the maturity paths and (3) for selecting KM instruments for those categories that have been neglected so far or (4) for selecting KM instruments that specifically target knowledge in incomplete maturity chains.

480. See also Figure B-24 on page 199.

Starting point for the maturity paths is person-product knowledge in the lower left corner of the classification diagram (I). The most important role is played by personal experience management which targets a particular type of knowledge of the least degree of maturity and thus is the starting point for a number of maturity paths. Knowledge systematically handled by individuals finds its way both into individual knowledge in routines (upper left corner, II) as well as into knowledge objects embedded in social systems (lower right corner, II). From there, knowledge finally enters the upper right corner (III) which contains those KM instruments that target comparably matured knowledge in organizations.

7.2.3 Size and media used

As opposed to e.g., relational data base systems, it is quite difficult to measure the size of the contents of KMS. In the case of relational data base systems, size is quite easily measured as the number of rows of a table times the number of bytes in every row. The sum total of all tables is the total size of a data base system. However, a “knowledge base” in most cases consists of a large number of knowledge elements, i.e. semi-structured files that are dispersed over a number of servers which not only contain files that are part of the KMS, but also more traditional documents which might also be managed with the help of a KMS.

Knowledge elements vary greatly in terms of *size* and in terms of *ICT used*, with respect to the type of ICT that is used to handle the knowledge elements, e.g., (relational) data base systems, word processing software, office information systems, file server, data warehouses, archiving systems, DMS, forums, Weblogs, Wikis or other CMS, Web server, video server, learning management systems, mailboxes or news server. Knowledge elements can also be organized in a variety of ways[481].

The *size of the knowledge base* is assessed using the following measures:

- the number of knowledge elements,
- the amount of storage capacity used (in MB).

The average size of knowledge elements will be calculated in order to get a more detailed picture about what an organization terms a knowledge element.

It is hypothesized that organizations with a systematic KM initiative store greater volumes of knowledge elements than organizations without one. In several related empirical studies, identification, providing access to and/or documentation of existing knowledge turned out to be among the first activities of KM projects[482]. The result of these activities should lead to more knowledge elements. These organizations should therefore use increased amounts of storage capacity for knowledge elements:

Hypothesis 15: Organizations with systematic KM handle a larger knowledge base than organizations without such an initiative

481. See section 7.2.4 - “Structuring of contents” on page 298.
482. See chapter 10 - “Related Empirical Studies” on page 439.

Also, organizations with KM initiatives are expected to handle a large amount of electronic resources that could be considered as knowledge elements with heterogeneous formats and types of media. The file format is not sufficient to determine the content or purpose of a knowledge element, e.g., an XML file can be technically a text processor document, a spreadsheet, or a scalable vector graphic (SVG), conceptually an idea, a lesson learned, a good practice or a skill description. KMS primarily deal with semi-structured, compound documents containing coded information for different purposes. However, the type of media has great impact on the requirements for meta-data management, e.g., a full text search may lead to a feasible result for a text document, but not for an image. The following *types of media* can be used in organizations[483]:

(Hyper-)text documents. Documents are stored in varying formats, e.g.:

- *document exchange formats:* such as the document exchange format rich text format and the formats developed by Adobe Systems postscript or the portable document format,
- *text document formats:* as part of office application suites, such as the Adobe Framemaker format, the Microsoft Word format or the Star/OpenOffice format,
- *hypertext documents:* e.g., Web pages, written in Hypertext Markup Language (HTML) or written in eXtensible Markup Language (XML). The latter can be characterized as a meta language which is used to integrate documents, data base outputs and various types of multimedia elements in a flexible way.

Multimedia contents. Multimedia contents could also be part of hypertext documents[484]:

- *audio files:* coded in formats, such as MPEG–Motion Picture Expert Group's MPEG Audio Layer III and the MP3 compression format, Dolby Laboratories Inc.'s format AC-3, Sun's Audio File format or Microsoft's WAVE format,
- *video files:* coded in different formats, such as the MPEG's format family of the same name, Real Network's RealMedia format or Microsoft's Audio-Video-Interleaved format,
- *vector graphs:* coded in formats like Computer Graphics Metafile CGM, Initial Graphics Exchange Standard IGES, AutoCAD's Drawing Exchange Format DWF/DXF or 3D-graphs, written in Virtual Reality Markup Language VRML,
- *pictures:* coded in formats such as the Bitmap format commonly known in the Windows world, the Graphics Interchange Format, the Tagged Image File Format TIFF, the UNIX graphic data format XPM and the compression format of the joint photographers expert group JPEG.

---

483. For a good overview of multimedia and electronic publishing formats see Steinmetz/Nahrstedt 1995, Henning 2000.

484. In the category (hyper-)text documents the focus is still on the text component whereas in multimedia contents the focus shifts to audio or video files, graphs or pictures. In the following, formats can be codecs, file layouts or both; see also Henning 2000 for details.

Contributions to newsgroups. These are regularly email (text) messages with or without attachments that are sent to discussion lists.

Data base elements. This type of media represents the traditional, structured form of data storage in hierarchical, network, object-oriented, multi-dimensional or, most commonly, relational data bases and data warehouses (for an overview of data base theory, development and systems see e.g., Elmasri/Navathe 1994, Inmon/Hackathorn 1994, Atzeni et al. 1999, Watson 1999). Data base elements still might be considered as part of a KMS's storage system, especially when connected to richer media like documents, multimedia contents and the interactive side of a KMS like contributions to newsgroups or email messages.

Organizations with a systematic KM initiative might also include more differing types of media in their knowledge bases than organizations without one. This should be especially true for multimedia elements, contributions to newsgroups and data base elements, whereas traditional documents could represent a smaller share of the knowledge base. Again, the activities identification, providing access to and/or documentation of existing knowledge should lead to a greater variety of types of media used to represent knowledge elements. Therefore, these organizations should use more variety in the types of media used:

Hypothesis 16: Organizations with systematic KM handle a higher share of multimedia elements, contributions to newsgroups and data base elements in their KMS than organizations without such an initiative

7.2.4 Structuring of contents

In addition to type of contents, the size and the media used in KMS, structuring and organizing the contents is supposed to be one of the key tasks in knowledge management. There have been many approaches suggested to organize knowledge in organizations that basically fall into two groups. On the one hand, AI methods are suggested to support the development of ontologies in organizations (e.g., Staab et al. 2001). On the other hand, business processes models are used as a starting point to identify the most critical business knowledge in organizations (e.g., Remus 2002). However, the interviews showed that in the organizations so far mostly pragmatic approaches are applied. In most cases, the knowledge structure is determined by a committee in a workshop without much methodical support and then evolves with new additions to the knowledge base. The investigation of knowledge structure will therefore be limited to a set of basic criteria to study to what extent organizations structure and organize their knowledge bases[485].

485. The interested reader will find a host of literature in the AI field that has a long tradition in dealing with structuring expert systems and knowledge bases and recently has been applied to broader domains, such as organizational document bases or Intranets (for links to literature on AI see also section 4.1.1 - "From organizational learning to knowledge management" on page 22).

The structure and organization of knowledge elements supposedly strongly influences the usefulness of a KMS. Structure not only determines how quick a participant can navigate to the knowledge elements needed, but also supposedly influences participants' mental models of the organizational knowledge base. Thus, structure and organization has a descriptive and a normative component influencing the way of thinking of the members of the organization. Structuring of contents will be assessed using the following two criteria:

- the number of knowledge clusters and the ratio between the number of knowledge clusters and the size of the knowledge base,
- the way of structuring: hierarchy, network or both.

According to the interviews, the hypertext is the single most important metaphor for organizing documents in an organizational Intranet or KMS. Navigation of hyperlinked documents has become a basic standard. The next step would then be to use the hypertext or network metaphor not only for navigation within documents, but also for the overall organization of knowledge areas. Thus, the network is supposedly the predominant principle of structuring knowledge areas when compared to the hierarchy.

Hypothesis 17: There are more organizations which apply a network structure to their knowledge areas than organizations with a hierarchical structure of knowledge areas

The interviews showed that organizations differ with respect to centralization of their KM tasks. It seems that organizations are facing a *trade off* between actuality/flexibility and understandability/simplicity of knowledge structure and contents. Actuality and flexibility of contents on the one hand require a decentralization of the corresponding KM tasks, e.g., storing of new knowledge, integration of knowledge in existing structure and especially update of structure. On the other hand, the more decentralized these tasks are, the more complex the contents might be due to the agglomeration of the variety of mental models held by the members of the organization that is not integrated.

However, as mentioned above it is a challenging task even for knowledge managers to determine the size and structuring of an organization's KMS. As a consequence, in the empirical study there will probably not be enough data on each of these measures to test correlations between complexity of contents and, say, a form of organizational design of the KM initiative or types of Groupware platforms and KMS used.

7.2.5 Quality of contents

The quality of contents is a key factor that determines the usability of a knowledge management system. Research on data and information quality has a long tradition in MIS and has been influenced strongly by quality management as well as knowledge management literature[486]. A large number of quality criteria have been sug-

486. Eppler 2003, 23, 41ff and the literature cited there.

gested that can be applied to measure or estimate the quality of contents of a KMS (Eppler 2003, 63).

Many authors have compiled lists of criteria to assess the quality of data[487]. Table B-17 shows a list of criteria that are widely used in the literature and in practice together with their description. However, the criteria for data quality are focussed on (raw) data, rather than on their interpretation by users and their combination, integration and contextualization. In order to be applicable for knowledge management, the criteria have to be extended and structured.

Eppler (2003) suggests a list of criteria for information quality together with their opposites (Table B-18).The criteria are structured according to the "level" of information quality and can be interpreted with respect to their application to content of KMS as follows:

- *infrastructure*: the infrastructure level deals with the quality of the knowledge management system that conveys the content.
- *process*: criteria on the process level help to evaluate knowledge processes and (parts of) knowledge-intensive business processes.
- *product*: the product level covers aspects of the resulting knowledge elements, i.e. the contents in a narrow understanding.
- *community*: finally, the community level deals with the knowledge receivers and covers the reconstruction process and the application of knowledge in the receivers' application domain and situation.

TABLE B-17. Criteria for data quality[a]

criterion	description
accuracy	data are precise enough for certain application areas
availability	data are available with respect to time and location of their user
completeness	all data are available that are needed for certain application areas
consistency	data correspond to the description in a repository; data are compatible with other data in the data base
correctness	data correspond to the portion of reality they describe
credibility	data can be traced back to a trusted source and transformations can be explained
relevance	data carry meaning for certain application areas
understandability	data are presented in a comprehensible form

a. Based on Schwinn et al. 1998, 210f.

These criteria are particularly important for documented knowledge elements stored in a KMS that are to be reused effectively and especially efficiently. Specific

487. For example Schwinn et al. 1998, 210f

functions and layers of KMS[488] contribute towards fulfilment of these criteria. Thus, the criteria for information quality can also be assigned to the layers of a KMS architecture so that each layer can be evaluated according to a number of specific criteria[489].

Eppler identified 28 "activities"[490] in a number of case studies that might increase the quality of contents (Eppler 2003, 82ff):

- *integration activities*: visualize concepts, list sources, summarize, personalize, prioritize contents, highlight aspects, give an overview, elicit patterns,
- *validation activities*: evaluate source, indicate level of certitude/reliability, describe rationale, compare sources, examine hidden interests/background, check consistency,
- *contextualization activities*: link content, state target groups, show purpose, describe background, relate to prior information, add meta-information, state limitations,
- *activation activities*: notify and alert, demonstrate steps, ask questions, use mnemonics, metaphors and storytelling, stress consequences, provide examples, offer interaction.

TABLE B-18. Criteria for information quality[a]

level	**criterion**	**opposite**
infrastructure level	accessibility	inaccessibility
	maintainability	neglect
	security	exposure
	speed	slowness
process level	convenience	inconvenience
	interactivity	rigidity
	timeliness	lateness
	traceability	indeterminacy
product level (soundness)	conciseness	polixity
	consistency	inconsistency
	correctness	falsity
	currency	obsolescence

488. See section 7.3.3 - "Integrating architectures for KMS" on page 311.
489. See section 7.8 - "Résumé" on page 390, particularly Table B-21, "Assignment of quality criteria to levels of KMS architecture," on page 391.
490. In the terminology of the activity theory, these "activities" might be considered as actions, i.e. routinized activities; see section 6.6.2 - "Activity modeling" on page 250.

TABLE B-18. Criteria for information quality[a]

level	criterion	opposite
community level (relevance)	accuracy	inaccuracy
	applicability	uselessness
	clarity	obscurity
	comprehensiveness	incompleteness

a. Source: Eppler 2003, 68.

These activities can be institutionalized in the form of e.g., the role of a subject matter specialist and the establishment of knowledge processes that are specifically designed to improve the quality of documented knowledge.

7.3 Architectures and services

Architectures in general play an important role in MIS as blueprints or reference models for corresponding implementations of information systems. The term architecture as used in MIS origins in the scientific discipline architecture and is used in a variety of ways, e.g., application architecture, system architecture, information system architecture and especially software architecture[491]. The prevalent architectural design recently has been impacted profoundly by the ideas marketed under the term service-oriented architecture (SOA). The primary concept of this architectural paradigm is discussed from the perspective of KM in section 7.3.1. Section 7.3.2 then reflects on some issues involved when designing a KM service infrastructure. Finally, section 7.3.3 reviews a number of theory-driven, vendor-specific and market-driven architectures of KMS and discusses their advantages and shortcomings.

7.3.1 Knowledge management service

Generally, a service is an abstract resource that represents a capability of performing tasks that form a coherent functionality from the point of view of providers entities and requesters entities (W3C 2004a, b). It consists of a contract, interfaces as well as implementation and has a distinctive functional meaning typically reflecting some high-level business concept covering data and business logic (Krafzig et al. 2005, 57-59). The service concept has gained much popularity with the advent of a set of standards that allow for open interaction between software applications using Web services[492]. A Web service is a software system, identified by a URI, whose public interfaces and bindings are defined and described using XML. Its definition can be discovered by other software systems. These systems may then interact with the Web service in a manner prescribed by its definition, using XML-based messages conveyed by Internet-based protocols (W3C 2004a),

491. See Lehner et al. 1995, 58ff for a definition and overview of the term.

see also (Alonso 2004, 124). Web services are one way of implementing business and technical services in a service-oriented architecture. A service-oriented architecture is based on the concepts of an application frontend, services, service repository and service bus (Krafzig et al., 2005, 57) which together make business and technical functions available as independent services that can be accessed without any information about their implementation.

The service concept has had a profound impact on enterprise application integration, on business-to-business applications and generally on the way information and communication infrastructures are designed and managed from a technical perspective (e.g., Cox/Kreger 2005). In addition to this technical impact, "SOA-enabled" businesses and organizations are sometimes called agile, on-demand or service-oriented enterprises, metaphors that attempt to carry over SOA semantics to organizational design (Bieberstein et al. 2005) which has connotations for changes in IT's general role in business (transforming business models), value creation (value networks), business processes (dynamically designed, net-like with emphasis on parallel processing) as well as organizational structure (service consumer-provider relationship complementing or even replacing traditional hierarchies; Cherbakov et al. 2005, 659). In the following, this section will concentrate on the specifics of the service concept applied to KMS (see also Maier/Remus 2007).

KM services are a subset of services offered in an organization, both basic and composed, whose functionality supports high-level KM instruments as part of on-demand KM initiatives. Examples for these services are *find expert*, *submit experience*, *publish skill profile*, *revisit learning resource* or *join community-of-interest*. Services are offered by service providers that procure the service implementations, supply their service descriptions, and provide the necessary support. Often, KM services cater to the special needs of one or a small number of organizational units, e.g., a process, a work group, a department, a subsidiary, a factory or an outlet in order to provide a solution to a defined business problem. KM services describe individual aspects of KM instruments implemented in heterogeneous application systems that can be combined into an enterprise knowledge infrastructure.

---

492. In distributed systems, service-oriented architectures can be seen as successors of component architectures. The underlying conceptual change could also trigger a paradigm shift from a primarily production-oriented view, not only of software production, to a view that takes into account the specifics of the service sector which has experienced growth during the last decades as opposed to the production sector which has declined. There is currently an initiative led by IBM and Oracle, but also involving institutions such as the European Commission, that aim at defining a research agenda for so-called services sciences. This agenda should bring the vision of a service-led economy to the focus of a number of scientific disciplines. Thus, the service concept transcends the scientific disciplines of computer science and information systems and also involves disciplines such as management, economics or service engineering.

7.3.2 Service infrastructure

Basic services can be composed into new composite services enabling larger integrated KM services. In addition, service descriptions have to be published in order to provide information about service capability, interface, behavior, and quality (Papazoglou/Georgakopoulos 2003). Figure B-51 shows the main layers of a KM service infrastructure.

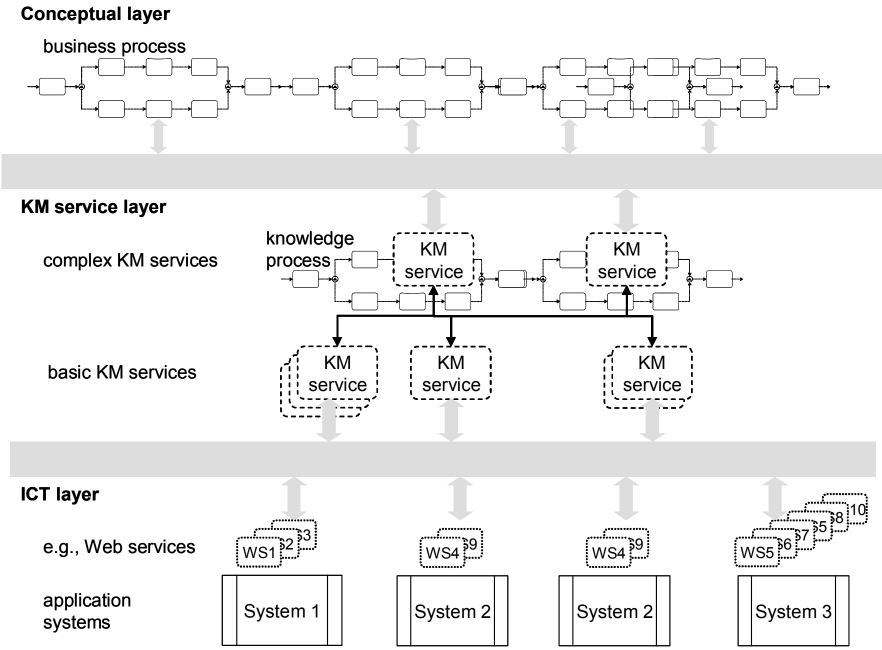

FIGURE B-51. KM service infrastructure[493]

Conceptual layer. Based on process descriptions, the conceptual layer defines which services are required in which core business processes, which services are offered by what service processes, who is responsible for them and what resources are allocated to fulfil them. Especially concepts of process-oriented KM can help to analyze, understand and design business and knowledge processes with regard to a knowledge-oriented and at the same time a strategic perspective on KM services in business processes.

ICT layer. Services are described, discovered and invoked with the help of negotiated or standardized sets of technologies, e.g., in the case of Web services WSDL, UDDI and SOAP. These technologies support the integration on different levels, i.e. human-to-machine, machine-to-machine and inter-organizational integration

493. Source: Maier/Remus 2007, 10

(Puschmann/Alt 2005). The ICT layer comprises infrastructure, integration, knowledge, personalization and access services dispersed over a variety of heterogeneous application systems that cover structured as well as semi- or unstructured data sources.

KM service layer. The main task is to bridge the gap between the conceptual and the ICT layer. KM services have to be composed using services offered by heterogeneous application systems from the ICT layer. In addition, discovery, call and provision of KM services from different activities of business processes have to be supported.

In the following, the conceptual layer is briefly reviewed[494]. Then, the primary function of the KM service layer is outlined with the help of an example. Finally, section 7.4 - "Centralized architecture" on page 318 presents the most important services that are required in order to implement a comprehensive KMS. These services, however, do not necessarily have to be implemented as one centralistic system, but can be accessed from different application systems using the service infrastructure described here.

Conceptual layer. The idea of a KM service infrastructure is demonstrated using a real-life example of a knowledge process and its composition by KM services. Identification, separation and description of relevant processes are important prerequisites. Models that support the conceptual layer were developed as part of a process-oriented KM modelling project[495]. In this project, a complex process landscape consisting of several knowledge processes was defined and modelled (Maier/Remus 2003). In extension to this project, the conceptual layer of a KM service infrastructure requires different levels of abstraction.

The highest level displays the activity and process landscape that shows the definition of processes as well as the assignment of KM instruments to KM activities. The second level refines the delineation of the processes that are shown in the first level e.g., by using event-driven process chains (Scheer 2001). The third level details these processes with the help of action charts linking single activities to knowledge structures. These models can be the first step towards the description of KM services together with their triggering events, inputs, outputs of activities and corresponding ICT systems and tools. In this project, modeling techniques provided by the ARIS (architecture of integrated information systems) method and toolset (Scheer 2001) were used. However, the development of a KM service infrastructure is not tied to a specific modeling technique as long as other methods provide techniques for modeling business processes on different levels of abstraction and a model type corresponding to action charts in ARIS[496].

494. For a detailed description see section 6.3 - "Process organization" on page 207.
495. The project is described in section 6.3.3 - "Example: Process-oriented KM" on page 217.
496. Examples for other relevant modeling approaches are mentioned in section 6.6.1 - "Process modeling" on page 240.

Action charts illustrate which service objects are consumed, produced and transformed. Here, these service objects are typically knowledge elements.

In general, service descriptions have to provide information about (Papazoglou/Georgakopoulos 2003):

- *service capability* states the conceptual purpose and expected result of the service by the description of output objects,
- *service interface* publishes the service's signature (input/output/error parameters and message types),
- *service behavior* can be described as detailed workflow invoking other services,
- *quality of service* publishes functional and non-functional quality attributes (e.g., service metering, costs, performance metrics, security attributes).

Figure B-52 shows the example knowledge process *knowledge documentation*, consisting of the two parallel sub-processes *content* and *skill management* with its main activities and triggering events. Processes were modelled as event-driven process chains (Scheer 2001).

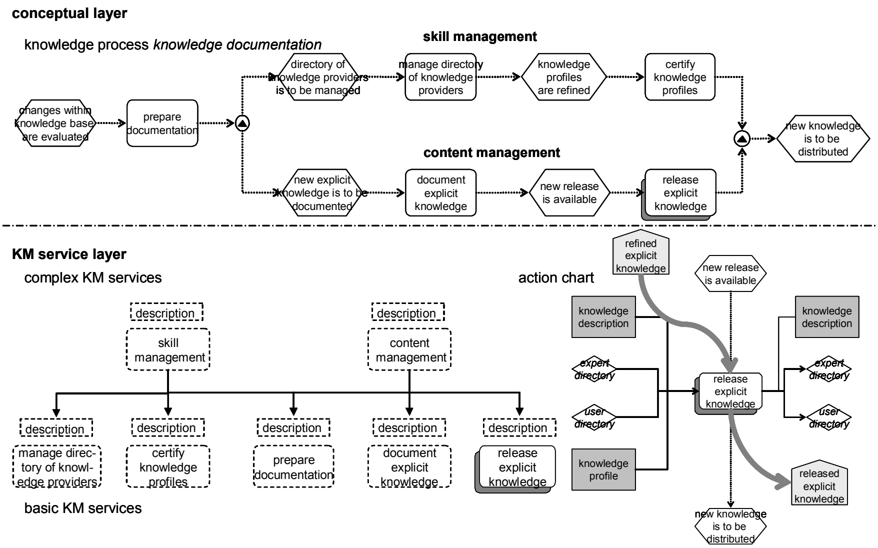

FIGURE B-52. KM services of the knowledge documentation process[497]

Every event-driven process chain is represented as a diagram. The recommended direction of reading is from left to right. Functions represent tasks or activities performed as part of the interactions from one or more objects. They are displayed as a rectangle with rounded corners. Functions produce events or states which in turn can cause a change of states of these objects or the execution of other

497. Source: Maier/Remus 2007, 12

functions. Events specify relevant states for objects that must be satisfied before functions can be executed and are displayed as hexagons. To display possible alternatives of similar business processes in one diagram, the event-driven process chain contains logical operators (OR, XOR, AND) that are used to describe the control flow between sequences of actions.

Experiences (i.e. lessons learned) that have been documented during the execution of business processes have to be managed regularly by initiating the process *knowledge documentation*. In order to avoid information overload and to guarantee a high quality standard of the knowledge base, changes within the knowledge base have to be evaluated. Therefore, appropriate measures to value, refine, certify and release knowledge have to be carried out (link to the process *enhancement of the knowledge base*). It is important to distinguish between explicit and implicit knowledge, since both types need different measures handling them. Explicit knowledge can be documented directly whereas implicit knowledge can be addressed by developing and maintaining an expert and user directory in which knowledge profiles are provided and linked to content in the knowledge base.

The result is an updated knowledge base with knowledge that can be used within business processes. It contains updated knowledge profiles of employees together with documented knowledge. Both are linked to functions in the business processes with the help of the process-oriented knowledge structure. A subject matter specialist can then release parts of the updated knowledge base for distribution. In addition, refined and updated knowledge profiles have to be certified, e.g., by discussions between supervisors or project managers and group or team members[498].

The next step is to determine which services are required to fulfil the process. At one extreme, the process can be viewed as one single, but complex service; at the other extreme, service granularity could be so fine that each function in the process can be constructed from multiple services. Similar to concepts in SOA, the choice is made by balancing quality of service characteristics (QoS), ubiquitous service reuse, and reduction of complexity for service composition (Crawford et al. 2005).

KM services can be viewed as encapsulated KM activities, accessible by an interface and described by action charts (providing an initial service description). The composition of KM services is presented in Figure B-52, together with one detailed service description (as action chart) for the KM service *release explicit knowledge* in the process *knowledge documentation*. This KM service approves content and makes it accessible to the employees of the organization. It releases knowledge descriptions, user and expert dictionaries, and assigns appropriate user privileges for the envisioned target group. It is based on the input *refined explicit knowledge* and produces the output *released explicit knowledge*.

ICT Layer. The ICT layer describes the services offered by heterogeneous application systems that have to be selected, called and combined in order to provide basic KM services. A comprehensive platform-type solution for these services has

498. See also section 6.2.2 - "Product-oriented instruments" on page 200.

been termed an enterprise knowledge infrastructure (Maier et al. 2005). From an ICT perspective, services can be structured into the following categories: (1) infrastructure services, (2) integration services, (3) knowledge services, (4) personalization services and (5) access services[499]. These categories help to structure existing services offered by different application systems. Next to semantic integration[500] between these services, process integration is required in the form of KM service composition which is explained in the following section.

KM Service layer. Regardless of the implementation, it is important to understand the steps required to decompose a process into a series of complex and basic services and operational characteristics (Crawford et al. 2005). Composing KM services means specifying how these services have to be discovered and selected (discovery), how they have to be accessed from different activities of business processes (call) and finally how these services are provided by the service infrastructure accessing heterogeneous application systems from the ICT layer (binding, provision). Modeling techniques help defining the composition of services (Crawford et al. 2005). Figure B-53 shows the interplay between conceptual and ICT layers by the example of invoking the complex KM service *search for experts* from the business process layer.

On the conceptual layer, this KM service has to be described using knowledge process descriptions and action charts specifying basic input and output parameters. Area of expertise is required as the minimum input parameter. Further input parameters can be specified that describe the context of the situation in which the service is invoked. Examples for context parameters are (1) process, i.e. the business process or task that the person is currently engaged in, (2) person, i.e. the profile of the person invoking the service, e.g., areas of expertise or skill levels, (3) preferences, e.g., for synchronous versus asynchronous communication channels, (4) products, i.e. electronic resources concerning the area of expertise that have been collected and/or analyzed by the person, e.g., learning resources, handbooks, reports or lessons learned, (5) applications and appliances, e.g., a Web browser on a desktop PC or a mobile application on a smartphone, (6) location, e.g., GPS coordinates or the connection, e.g., wired LAN, wireless LAN or UMTS connection, (7) date and time, normalized according to the time zone, which might help to determine the appropriate way of contacting experts and (8) urgency of the need for an expert. Execution of the service results in a list of experts, brief descriptions, contact history and information about the (social) relationship to the searcher, e.g., common business acquaintances, and contact and availability details, ordered according to the preferences of the experts together with links to further KM services that can be invoked in order to establish a connection to the selected expert.

The middle layer in Figure B-53 shows the composition of a number of basic KM services into one complex KM service and maps the required basic KM ser-

499. See section 7.4 - “Centralized architecture” on page 318.
500. See section 7.7 - “Semantic integration” on page 374.

vices to actual, "real" services offered by diverse application systems that are part of the ICT layer. It is a structural abstraction of the composition process, while the interplay between the basic services can be described e.g., with UML activity diagrams or state-charts or with BPEL-oriented notations such as BPMN[501].

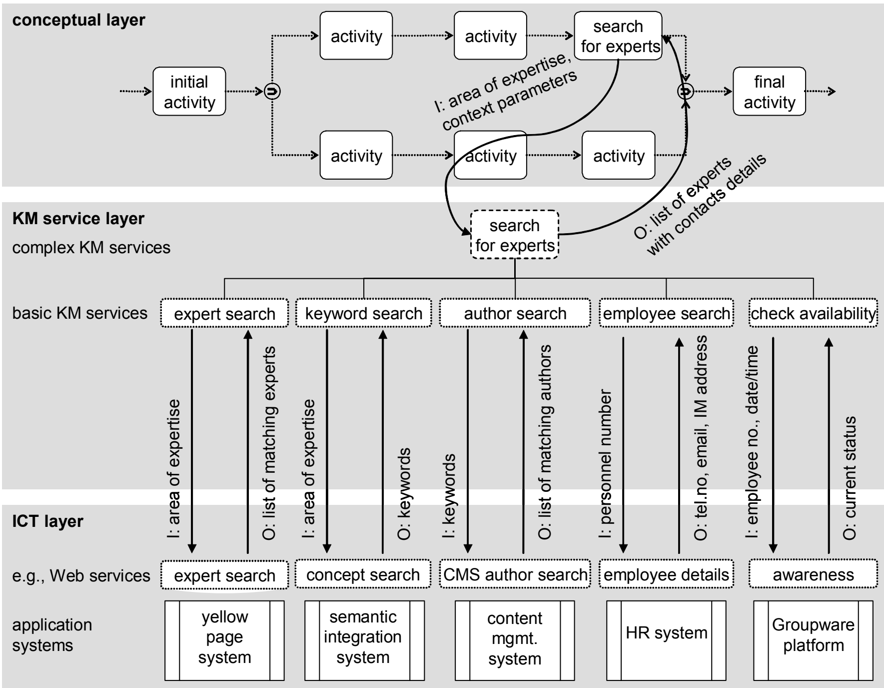

FIGURE B-53. KM service invocation

The complex KM service *search for experts* is composed of the basic KM services (1) *expert search*, (2) *keyword search*, (3) *author search*, (4) *employee search* and (5) *check availability*. The *expert search* service delivers a list of IDs (e.g., personnel numbers) for experts matching the input parameter of an area of expertise. The *author search* service requires a list of keywords describing the area of expertise. Thus, the complex KM service *search for experts* also comprises an integration function or invokes an integration service for the task of finding keywords that describe the area of expertise, here called *keyword search*. The keywords are assigned to areas of expertise either in a simple data base solution or in a more advanced semantic integration system based on an ontology. With the help of an inference engine, these relationships together with rules in the ontology can be used to determine a list of keywords[502]. The *author search* service then returns a

501. See http://www.BPMI.org for a description of the BPMI stack for the composition of Web services.

list of IDs of matching authors or active contributors to the CMS respectively. An *employee search* service takes the personnel numbers found in the *expert search* and the *author search* and returns contact details, e.g., telephone number, email address, instant messaging address. Finally, the *check availability* service delivers the current status of the experts and a decision on their availability.

The ICT layer binds the basic KM services of the conceptual layer to application systems in the current work environment of the searcher that can deliver these services. In the case example, there might be a yellow page system, a semantic integration system, two content management systems, an HR system, a Groupware platform and an instant messaging system that offer Web services fitting to the descriptions of the basic KM services on the conceptual layer. Depending on which systems are accessible to the calling complex KM service, the actual implementation could consist e.g., of basic services (1) and (4), of (3) and (4), of (1), (4) and (5), of (1), (2), (3) and (4) or of all services respectively. Consequently, the description of the complex KM service needs to include some specification of what basic KM services are mandatory and of what combinations of basic KM services are allowed. Figure B-53 shows the three layers and an example of calls of KM services from activities in business processes and their binding to the corresponding Web services on the ICT layer.

The KM service infrastructure supports service-oriented, agile or on-demand KM approaches in organizations that take into account decentral developments of KM initiatives. Thus, KM technologies have to operate increasingly on infrastructures that support the rapid deployment of relevant tools and systems for ad-hoc, intensive and inter-organizational collaborations (Tsui 2005). Recently, these dynamic approaches of bringing the right knowledge rapidly to the point where it is needed have been called just-in-time KM (Davenport/Glaser 2002), workplace learning (Ellström 2002, Illeris 2003) or on-demand KM (Sampson et al. 2002).

When designing and implementing KM infrastructures, KM initiatives can introduce service-orientation as additional guideline. The three-layered KM service infrastructure composes services from heterogeneous applications into specific KM services and supports their discovery, call and provision from activities within business processes. This infrastructure aims at solving the following challenges:

Strategy. Strategic alignment is realized by connecting KM services to the materialization of strategic decisions (e.g., customer orientation) in the form of business processes and corresponding application systems on the ICT level. The deployment of KM services in organizations might profit substantially from both, the integration and the corresponding alignment with strategic goals.

Processes. Process orientation is realized by not only focussing on business processes as main drivers for calling KM services, but also on knowledge processes

502. See section 7.7 - “Semantic integration” on page 374.

which comprise a procedural view of a bundle of KM instruments implemented by KM services that are in turn described with the help of action charts.

Instruments. The numerous KM measures, procedures, instruments or tools applied in isolation from each other are integrated by bundling KM instruments to provide complex KM services. Business processes determine which KM services are required in which core business processes, are offered by what service processes, who is responsible for them and what resources are allocated to fulfil them.

Architecture. A concise KM architecture consisting of a KM service infrastructure on different levels helps reducing complexity and improving flexibility of KM initiatives. One of the major advantages of a KM service infrastructure is the ability to build it once and reuse it frequently. However, the efforts to implement a KM service infrastructure should not be underestimated. Already established KM services have to be identified and made available. New KM services have to be implemented. KM services have to be composed and decomposed finding the appropriate level of detail. The quality of KM services has to be assessed and documented in order to provide a constant level of quality throughout the knowledge life cycle.

The KM service infrastructure can be considered as an approach of a strategy-based integration of KM services which provides a blueprint, i.e. a framework and platform for dispersed KM services defined in heterogeneous KM initiatives. In the following, the services required for comprehensive KM initiatives are structured according to an ideal architecture and then described in detail[503]. This can serve as a framework guiding the design of a KM service infrastructure and the integration of application systems towards a transparent, centralistic KMS solution.

7.3.3 Integrating architectures for KMS

There are basically three main sources for architectures describing the structure of knowledge management systems: theory-driven, vendor-specific and market-driven architectures which will be discussed in the following.

Theory-driven architectures. The first group of KMS architectures is the result of theoretic investigations which represent a theory-driven decomposition of an organizational knowledge base or organizational memory and derive ideal groups of functions or components of a corresponding ICT system respectively[504].

Core functions of KMS can be viewed and categorized on different levels and from different perspectives. Classifications of functions as found in the literature fall roughly into one of the following classes[505]:

503. See section 7.4 - “Centralized architecture” on page 318.
504. See for example Stein/Zwass 1995, 98; see also section 4.3 - “Knowledge management systems” on page 82.
505. For classifications of KMS see section 7.6.2 - “Classes” on page 369.

- **categorization on the technical level:**
  These are specific system functions, like workflow management functions, document management functions, communication functions etc.
- **system-centered categorization:**
  An example for a system-centered categorization is the distinction between integrative and interactive KMS. This perspective bundles functions into function areas which give an indication of the primary direction of the use of such systems. They are usually a combination of functions on the technical level.
- **categorization according to knowledge (management) tasks:**
  This can either be concrete phases of a knowledge life cycle like knowledge identification, acquisition, storing, distribution etc. or abstract "processes" such as externalization, internalization, combination and socialization (Nonaka 1991, 98f, Nonaka 1994, 18f).

A classification of KMS functions can focus on the system-centered categorization which is more abstract than the list of functions on the technical level and more specific to KMS than the classifications with respect to KM theory. The following list of function areas was derived (a) from an extensive survey of existing KMS (Maier/Klosa 1999c), (b) from a set of empirical studies on KM[506] and (c) from several approaches to classify functions of KMS in the literature[507]:

- knowledge search,
- knowledge presentation,
- knowledge publication, structuring and linking,
- knowledge acquisition,
- knowledge communication and cooperation,
- computer-based training and tele-learning,
- administration of KMS.

The function areas can be further aggregated. *Knowledge search* and *presentation* are both discovery-oriented groups of functions of KMS. Thus, they are two sides of the same coin and can be drawn together. *Knowledge publication, structuring and linking* as well as *knowledge acquisition* are oriented towards (structured) publication of knowledge elements and thus can be combined as well. An architecture for a KMS has to show how these function areas are realized.

Zack classifies KM tools and systems into one of the following two segments: KMS with an integrative versus an interactive architecture (Zack 1999a, 50). This classification corresponds to the two main directions of KM research, human orientation and technology orientation, and Hansen et al.'s (1999) distinction of KM strategies into a personalization versus a codification strategy[508]:

---

506. See chapter 10 - "Related Empirical Studies" on page 439.
507. Ruggles 1997a, 5ff, Angus/Patel 1998, Apostolou/Mentzas 1998, 3.3ff, Borghoff/Pareschi 1998a, 5ff, Warschat et al. 1999, 56f, Krallmann et al. 2000, 233f, Seifried/Eppler 2000, 29.

- **integrative knowledge management architecture:**
  Integrative KM applications focus a repository and the explicit knowledge it contains as the primary medium for knowledge exchange. Integrative knowledge management applications can be further segmented according to the extent to which knowledge producers and consumers come from the same knowledge community. On the one extreme (called *electronic publishing*), there is neither direct interaction nor a shared context (in terms of e.g., belonging to the same community of practice) between producers and consumers of knowledge. Consumers do not give feedback and do not modify the knowledge in the repository. On the other extreme (called *integrated knowledge base*), producers and consumers share context intensively (e.g., they belong to the same organizational unit or community of practice).
- **interactive knowledge management architecture:**
  Interactive KM applications primarily facilitate the exchange of tacit knowledge among people. If there is a repository, it is seen more as a by-product the content of which changes dynamically. Interactive KM applications can be further segmented according to the expertise level of producers and consumers and the degree of formalization imposed on the interaction. On the one extreme, there is a formal, well defined knowledge transfer between "knowers" and "not knowers" (called *distributed learning*). On the other extreme, there is ad-hoc or emergent interaction more like an electronic discussion space (called *forum*).

Holistic KMS implementations aim at bridging the gap between these two architectures, at their combination and integration into a single KMS architecture. As mentioned by Zack, context plays the key role in bringing these two architectures together. Apitz et al. (2002, 33) present a KMS architecture that emphasizes contextualization as an important cornerstone in KM (see Figure B-54). Context management handles the context of topics or themes, and the context of tasks and processes and is used (1) to support workflows, (2) to describe information sources and organizational knowledge, (3) to acquire information and (4) to refine information that is pushed to or pulled by the knowledge worker. Also, the architecture stresses the importance of an integration of an intelligent handling of information technologies (information sources) on the one hand and of knowledge-based technologies (knowledge base) on the other hand. In this architecture, "intelligent" means the consideration of the types of context for both, information sources and the knowledge base. However, the proposed theory-driven architectures are limited to the conceptual level and do not indulge into the depths of concrete ICT components, tools or systems.

---

508. See also sections 4.1.4 - "Definition" on page 52 and 5.2.3 - "Generic knowledge management strategies" on page 129.

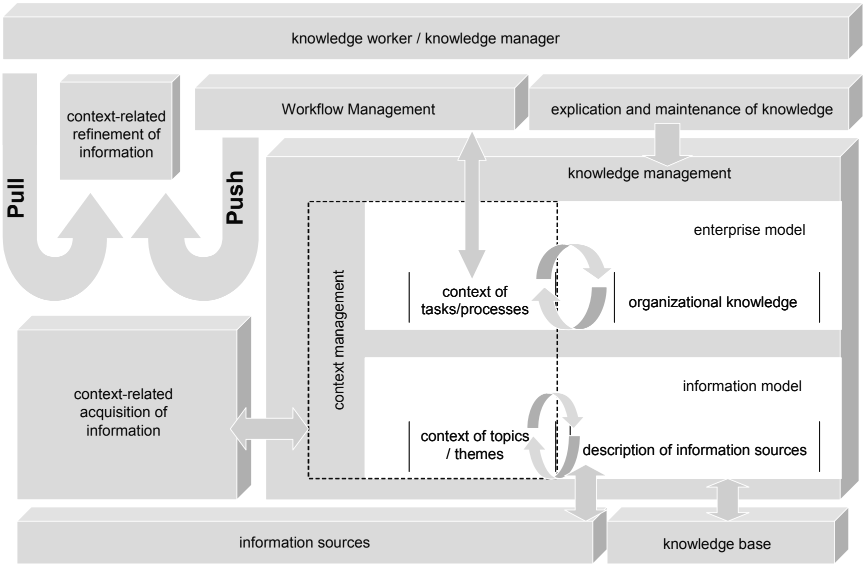

FIGURE B-54. Task-oriented architecture for KMS[509]

Vendor-specific architectures. Vendors of KMS publish white papers describing their perspective on knowledge management and place their tools in a KM architecture that regularly pays attention to the ICT infrastructure already available in the organizations[510].

Figure B-55 shows the simplest form of such an architecture. The KMS is just moved in between a standard Web browser and relevant data and document sources that exist in an organization. This approach is the traditional middleware approach that can be found in many KMS implementations.

Comprehensive KM suites comprise an often large number of modules offering functions such as text mining, tools for semantic integration of meta-data on data and documents, a search engine, visualization, administration of users and privileges, publishing and reporting.

509. Source: Apitz et al. 2002, 33.
510. See e.g., Baubin/Wirtz 1996, 139 for Accenture's Knowledge XChange, see Sippach et al. 1999, 65f for Multimedia Software GmbH's Intranet Knowledge Management System; see also the white papers on the homepages of KMS vendors: e.g., of the Empolis Knowledge Management Suite (Empolis), Hummingbird KM suite (Hummingbird, now Open Text), Hyperwave Information Server (Hyperwave), Intraspect 4 (Intraspect) or Livelink (Open Text). More recently, vendors modularize their offerings and package these modules according to application scenarios or concrete business needs for which the platform is used. Knowledge management is one of those application scenarios. For a more detailed analysis of Open Text Livelink see section 7.4.9 - "Example: Open Text Livelink" on page 336.

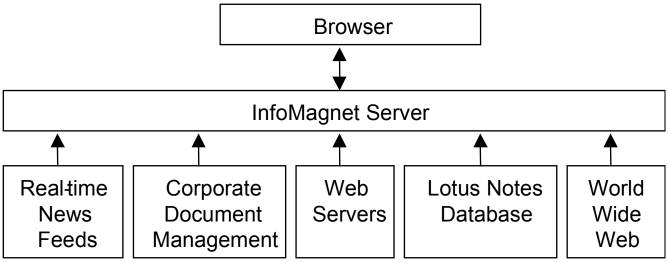

FIGURE B-55. Simple architecture for KMS[511]

Market-driven architectures. A third group of authors applies a more pragmatic approach and empirically distills the most important components of an organizational knowledge management environment which is integrated with more traditional data and document management systems as well as communication systems[512]. The authors mostly rely on the offers of (a number of) vendors of standard software tools, platforms and systems to support KM or analyze the individual KM environments of organizations that are regarded as KM pioneers and develop their own KMS solutions. These architectures are mostly layer models. The number, naming and inclusion criteria of the layers differ from author to author. Examples for layered KMS architectures are:

- OVUM, also a vendor of KMS tools, developed a simple architecture for KMS based on an empirical study on ICT demands and supplies for KM (Versteegen 1999). The architecture supports the four KM core processes capture, classification, sharing and understanding of knowledge and consists of six layers: (1) information and knowledge sources, e.g., texts, DBMS, email directories, WWW and the knowledge workers themselves, (2) infrastructure, i.e. email, file server and Intranet-/Internet-services, (3) information and process management that is located in a knowledge repository, (4) a shared taxonomy, a knowledge map, (5) knowledge management services for discovery and collaboration and (6) a user interface that consists of a knowledge portal.
- The architecture presented by Applehans et al. (1999) is quite similar to the OVUM architecture and also comprises six layers (see Figure B-56): (1) information and knowledge sources, called repositories, (2) transport layer, which corresponds to an Intranet infrastructure, extended by collaboration and stream-

511. Source: CompassWare 1998.
512. See e.g., Applehans et al. 1999, 87ff for a layered knowledge architecture, Bach 1999, 69 who proposes a tool architecture for business knowledge management, CZ 1999, 13 for the comprehensive KM architecture proposed by the Meta Group, Versteegen 1999, 118 who describes OVUM's six-layer KMS architecture, Seifried/Eppler 2000, 31ff who suggest a structured set of functionality expected from knowledge management suites and Vieser 2000 who presents the Siemens three-layered architecture for ICT tools, services and KM applications.

ing media tools, (3) application layer, with the examples calendar, yellow pages and analysis tools, (4) intelligence layer which consists of search, personalization and agent technologies, (5) access layer that stresses security technologies and (6) user interface, here as in most KMS architectures mainly a Web browser.

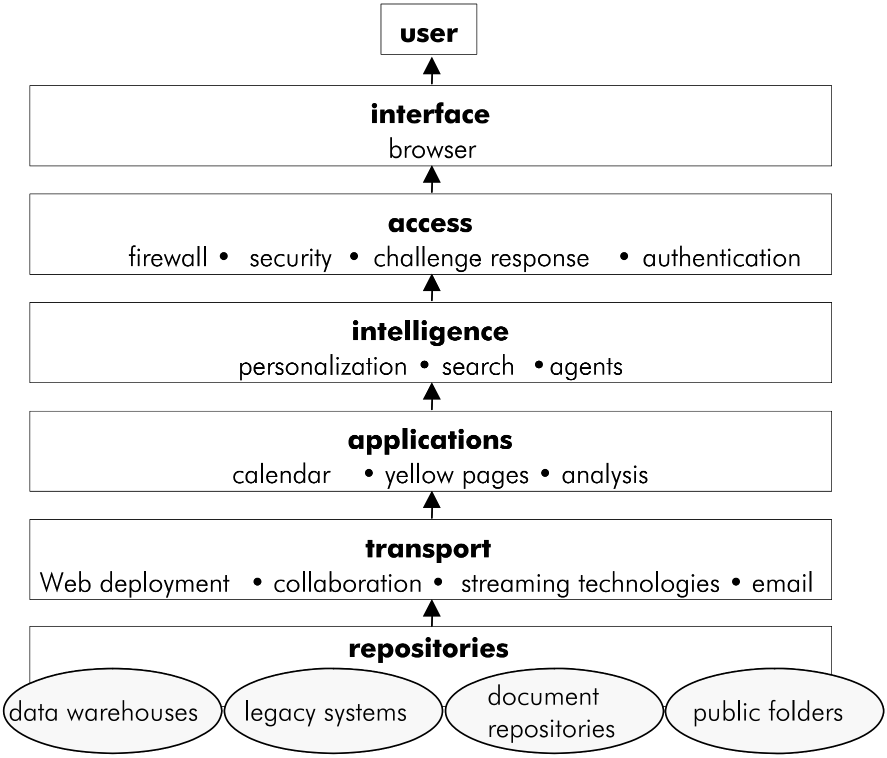

FIGURE B-56. Layered architecture for KMS[513]

- Bach's (1999, 69) architecture stresses the importance of supporting individual knowledge workers with an integrated electronic work place on the basis of a process-oriented knowledge management architecture (see Figure B-57). The architecture consists of the five layers (1) *Intranet infrastructure*, (2) a wide array of *information sources* that also contain transaction processing systems, data bases and external sources, (3) *integration services*, that contain a search engine, a data warehouse, a directory and a viewer for heterogeneous types of documents, (4) *information services*, that provide support for publishing, workflows, a library, an employee directory and collaboration tools and (5) the *integrated work place*. Bach's aim of integrating process-orientation into KMS

513. Source: Applehans et al. 1999.

architectures manifests on all layers except the integration layer. Task flows in the integrated work place are supported by workflows as part of information services. Transaction systems execute processes in the layer of information sources.

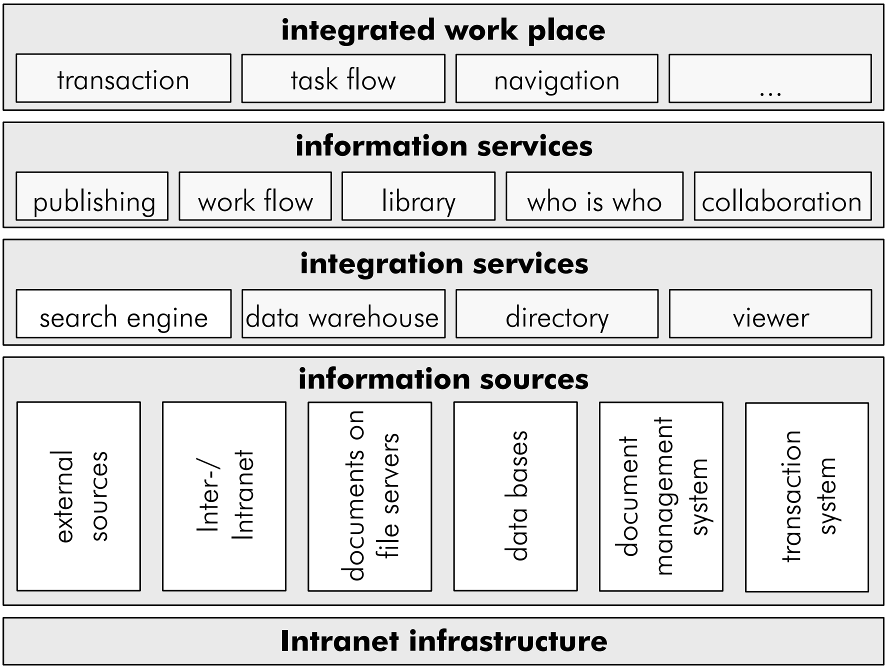

FIGURE B-57. Integrated, layered architecture for KMS[514]

- Becker et al. (2002, 24) present an architecture that views KMS from the perspective of meta-data (see Figure B-58). The starting point for this architecture is the observation that in most organizations there are already a number of application systems installed that provide a substantial portion of the functions that are required for KM. Becker et al. conclude that the KMS additions are basically restricted to the integration of these application systems with the help of a defined set of meta-data and a knowledge portal. Examples for application systems that provide KM functionality are content management systems, data warehouses, enterprise resource planning systems and workflow management systems. Each of these systems handles its own meta-data. Consequently, a KMS needs to align the meta-data of these systems. A knowledge management portal accesses the contents of these application systems on the basis of a separate data base with integrated meta-data drawn from these systems.

514. Source: Bach 1999, 69.

The comparison of these architectures reveals that each architecture suggests the establishment of a number of components organized on a number of layers, but none of the architectures comprises all the layers.

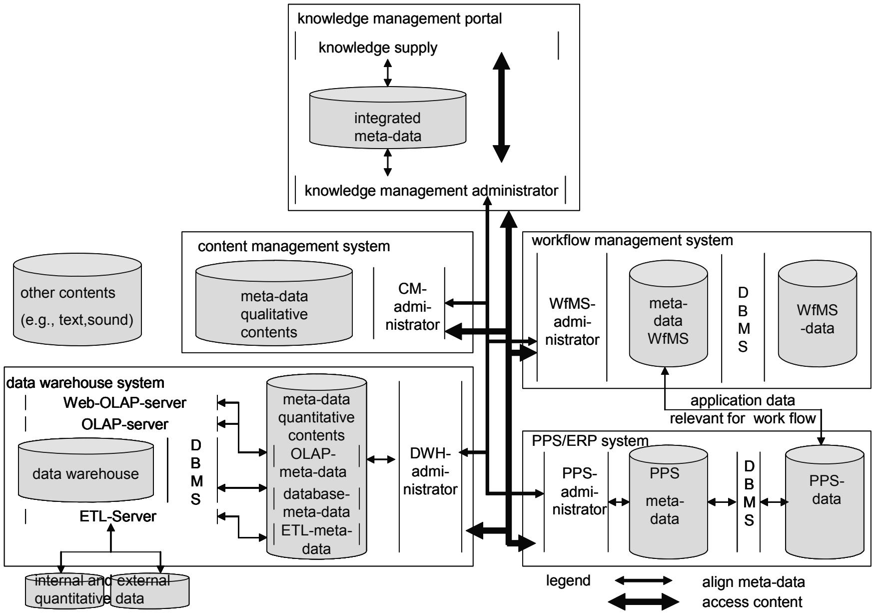

FIGURE B-58. Meta-data-oriented architecture for KMS[515]

For example, the OVUM architecture lacks a security layer, Applehans et al.'s architecture has no integration layer with a shared taxonomy and a repository. Bach's architecture provides the important layer of an integrated knowledge work place. However, the underlying layers lack detailing. Becker et al. finally introduce the aspect of a meta-data-based integration of legacy systems into a useful KMS. However, the role of KMS in this architecture is reduced to a portal. It lacks the intelligent functions that all other architectures stress as being one of the key components that distinguish KMS from traditional approaches.

7.4 Centralized architecture

The architectures described in the last section are now integrated into an ideal architecture for centralized KMS. Section 7.4.1 gives an overview of this architecture which comprises infrastructure, integration, discovery, publication, collaboration, learning, personalization and access services. In the sections 7.4.2 to 7.4.8, a comprehensive list of individual KMS functions are discussed structured according

515. Source: Becker et al. 2002, 24.

to the services organized in the architecture. Finally, section 7.4.9 discusses the components of Open Text Livelink according to the ideal architecture presented in section 7.4.1 as an example for a typical centralized KMS that is one of the best known and most widely used standard KMS in the market.

7.4.1 Overview

Figure B-59 shows an ideal layered architecture for centralized KMS that represents an amalgamation of the theory-driven, market-oriented and several vendor-specific architectures and integrates their components and layers.

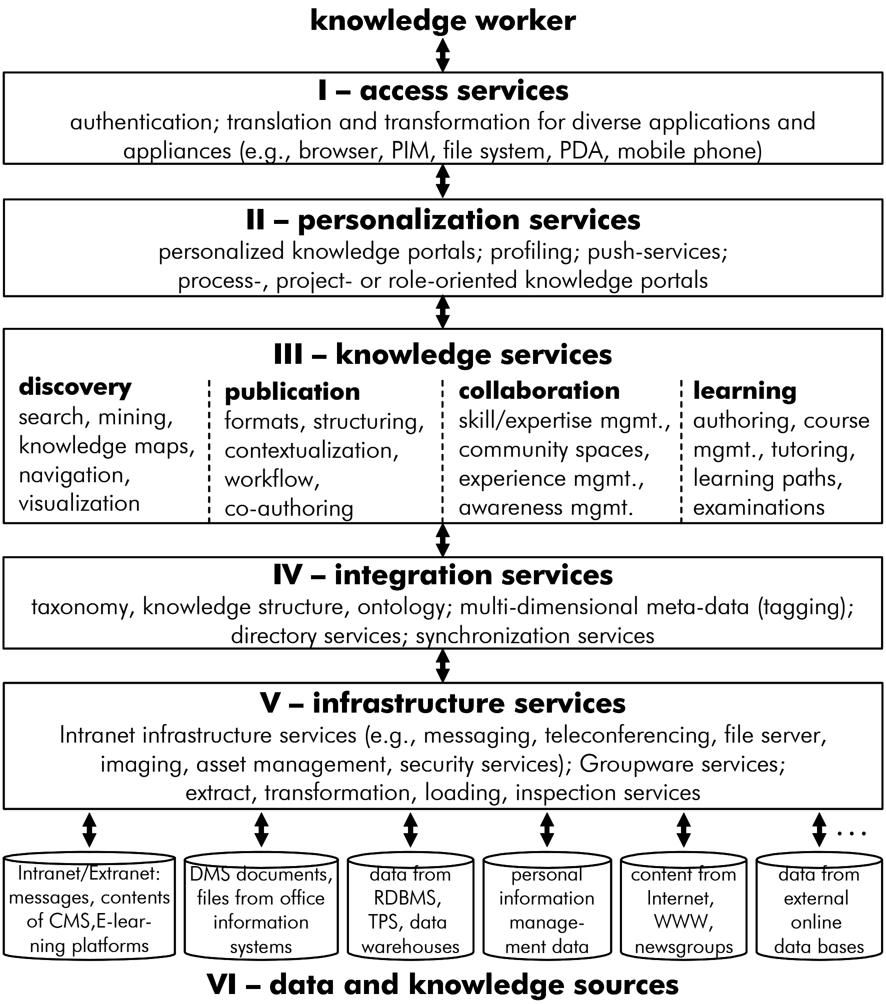

FIGURE B-59. Architecture for centralized KMS

As in the majority of architectural approaches in the literature, the ideal architecture is oriented towards the metaphor of a central KM server that manages all shared knowledge and can be accessed by knowledge workers, the clients.

Access services. The participant or knowledge worker accesses the organization's KMS with the help of a variety of access services, that translate and transform the contents and communication to and from the KMS to heterogeneous applications and appliances. Synchronization between these different applications and appliances including an integrated management of meta-data is provided by an integrated knowledge workspace. The KMS has to be protected against eavesdropping and unauthorized use by tools for authentication, authorization and encryption.

Personalization services. Main aim of the personalization services is to provide a more effective access to the large amounts of knowledge elements and thus to avoid information overload (Eppler/Mengis 2003). On the one hand, subject matter specialists or managers of knowledge processes can organize a portion of the KMS contents and services for specific roles or develop role-oriented push services. In this case, the knowledge services are accessed by the knowledge workers through an enterprise, a work group, a project or a role-specific knowledge portal respectively. On the other hand, both, the portal and the services can be personalized with the help of e.g., interest profiles, personal category nets and personalizable portals. Automated profiling can aid personalization of functions, contents and services.

Knowledge services. The core knowledge processes—search and retrieval, publication, collaboration and learning—are supported by *knowledge services.* These are key components of the KMS architecture and provide intelligent functions for:

publication: is the joint authoring, structuring, contextualization and release of knowledge elements supported by workflows,

discovery: means search, retrieval and presentation of knowledge elements and experts with the help of e.g., search, mining, visualization, mapping and navigation tools,

collaboration: supports the joint creation, sharing and application of knowledge of knowledge providers and seekers with the help of e.g., contextualized communication and coordination tools, location and awareness management tools, community homespaces and experience management tools and

learning: is supported e.g., by authoring tools and tools managing courses, tutoring, learning paths and examinations.

Integration services. Knowledge services work on the basis of integration services, e.g., a knowledge repository which handles the organization's meta-knowledge describing knowledge elements that come from a variety of sources with the help of meta-data for a number of dimensions, e.g., person, time, topic, location, process, type[516]. A taxonomy, a knowledge structure or an ontology help to meaningfully organize and link the knowledge elements and are used to analyze the semantics of the organizational knowledge base. Moreover, integration services are needed to manage meta-data about the knowledge workers that work with the KMS, e.g., in directory services. Finally, synchronization services export a portion of the knowledge workspace for work offline and (re-)integrate the work on knowledge elements that has been done offline.

Infrastructure services. The personalization, knowledge and integration services layers together can be viewed as a KMS in a narrow sense[517]. These layers are based on an *Intranet infrastructure* which provides basic functionality for synchronous and asynchronous communication, the sharing of data and documents as well as the management of electronic assets in general and of Web content in particular. In analogy to data warehousing, extract, transformation and loading tools provide access to the data and knowledge sources[518]. Furthermore, inspection services (viewer) are required for heterogeneous data and document formats. Inspection services support viewing of documents without the corresponding application, e.g., a text document without the text processing software that created the document.

Data and knowledge sources. The data and knowledge source layer gives some examples of the wide variety of electronic sources for data and knowledge which have to be integrated into the KMS or at least accessed through the KMS. In addition to *organization-internal sources*, such as the organization's transaction processing systems, data base systems, data warehouses, document management systems, content management systems, messaging systems and personal (or group) information management systems, many organizations need to include *organization-external sources* into their KMS. There is a huge and growing market for external (on-line) data bases. They can be classified e.g., into *fact data bases* that contain large collections of data and *reference data bases* which collect literature and/or references to literature. Examples for some well-known data supply companies that operate hundreds of data bases are (Mertens/Griese 2002, 20ff): DIALOG Information Services, Lockheed Information System, Predicasts, Reuters, or the Statistisches Bundesamt in Germany. Last but not least, the Internet, especially the WWW and newsgroups, provide abundant material that has to be considered in a KMS architecture.

In the following, the functions of a KMS that are required to perform these services are discussed according to the layers shown in Figure B-59. The layers comprising KMS in a narrow sense - personalization, knowledge services and integration, are discussed in detail. Due to their importance as key components of KMS, the four bundles discovery, publication, collaboration and learning that together provide the knowledge services are discussed separately.

---

516. For a description of these dimensions see section 7.5.3 - “Example: Infotop” on page 349.
517. See section 4.3.2 - “Definition” on page 86.
518. The input part of a data warehouse architecture has been called data acquisition layer, Gray/Watson 1998, 17 or input layer, Muksch/Behme 1998a, 45.

7.4.2 Infrastructure and integration services

In addition to the publication of knowledge elements by participants, KMS should provide functions to transfer knowledge elements from sources external to the KMS into the system. The functions can support both, the manual and the automatic integration of knowledge elements from organization-internal and organization-external sources. Knowledge acquisition also comprises the value-added process of deriving knowledge (in the sense of interesting relationships, patterns) from large collections of data (data bases, data warehouses)[519].

- *manual integration of external knowledge elements:* this function supports the integration of e.g., documents, bookmarks, links, multimedia and hypermedia elements such as video files, audio files, graphics, pictures or integrated video, audio and text files into the KMS,
- *automatic integration of knowledge elements from external sources*: the KMS automatically searches a predefined domain of organization-external knowledge sources (e.g., with the help of intelligent agents, crawler) and integrates new or updated knowledge elements,
- *generation of knowledge elements from internal data sources*: this function generates reports from organization-internal data bases (e.g., production, sales or financial data) through a value-added process (i.e., advanced reporting functions),
- *statistical data analysis*: comprises techniques and functions that have been developed under the label *business intelligence* to help managers and analysts to discover relationships in large collections of data, e.g., data mining, knowledge discovery in data bases, on-line analytical processing, decision support systems as well as statistics software packages such as SAS or SPSS.

7.4.3 Discovery services

Functions for knowledge search provide together with functions for knowledge presentation the output-oriented part of a KMS and can be divided into pull-functions and push-functions (Horstmann/Timm 1998, 242f). “Pull” means that the participant (inter-)actively uses search (support) functions, such as keyword search, a thesaurus or navigation tools to retrieve knowledge elements. Push-functions are activated once as an information subscription, the start of an intelligent agent or an email to a listserver[520] and then deliver knowledge elements automatically when the function detects new and/or interesting knowledge elements within a certain period of time.

1. Primary search functions:

- *keyword search*: a widely used function with which keywords belonging to certain categories (e.g., author, title, year) are used to search for e.g., docu-

519. This process requires intelligent knowledge services that can also be applied to discover relationships between knowledge elements of the KMS’ own organizational knowledge base.

520. See also section 7.4.5 - “Collaboration services” on page 327.

ments or persons. The keywords can be combined using logical operators (Boole's algebra, e.g., AND, OR, NOT),

- *meta-search system*: is also called a multi-search system (Horn 1999, 59) and provides functions supporting the (user-friendly) access to multiple knowledge sources. The term "meta" denotes here that the meta-search system accesses several individual search systems and "forwards" the search term in order to provide search results that span several data or document bases. Meta-search systems are further distinguished with respect to the search domain which they support, such as organization-internal and/or organization-external systems and with respect to the formats which they can search[521]. Meta-search engines offered on the WWW (e.g., MetaGer, MetaCrawler, ProFusion) so far are limited to HTML pages[522],
- *user-initiated filters*: allow to restrict the search to e.g., certain knowledge sources, topics, time, formats to avoid or at least decrease irrelevant search results,
- *navigation*: instead of directly typing in keywords to search for knowledge elements, participants can navigate through the knowledge structure to find their way to knowledge elements. The knowledge structure can be presented using two- and three-dimensional (e.g., hyperbolic) visualization of categories (category browsing) as well as sitemaps to avoid the "lost-in-cyberspace" phenomenon. Navigation also comprises e.g., Web browser functions, such as going back, history or hyperlinks to related hypertexts[523].

2. Search support functions:

Search support functions are not search functions on their own, but can be applied so that the quality of search results is improved:

- *thesaurus/synonyms*: a thesaurus is an alphabetically or otherwise systematically organized directory of words which displays the terminological relationships between the words (e.g., homonyms, synonyms) within a certain domain (Mertens et al. 1997, 408f) A thesaurus can either be used intentionally or automatically by the KMS to improve participants' search terms,
- *presentation of new/unread documents*: knowledge elements which have been added to the KMS that is searched (unspecific novelty) and/or which have not been accessed by the participant (participant-specific novelty) are marked, e.g., by a symbol or by using a specific color,
- *search assistants/search support*: aid the participants during the search, e.g., on-line help, tips, context-specific help, e.g., help in narrowing or extending a search term,

---

521. Examples are documents, hypertext, relational data bases, data warehouses and the like; see also section 7.2 - "Contents" on page 281.
522. See also section 7.1 - "Technological roots" on page 273.
523. See also section 3. - "Presentation of relationships between knowledge elements before search:" on page 324.

- *display of access statistics for knowledge elements*: the KMS displays the numbers of accesses to certain knowledge elements and/or knowledge areas. The participant can use this information, e.g., to get a feeling of how many other participants have been interested in a topic or to detect trends.

After searching and navigating the knowledge space, knowledge presentation comprises functions that support the presentation of search results and that visualize the organization of knowledge elements, their structure and the relationships between knowledge elements.

3. **Presentation of relationships between knowledge elements before search:**
   - *three-dimensional visualization*: this function uses three-dimensional models to represent the organization of knowledge elements and their relationships. *Hyperbolic browsers* use mathematical models to visualize three-dimensional hierarchic structures on a two-dimensional medium (screen, paper). Examples are the tools PersonalBrain (TheBrain Technologies) and InXight SmartDiscovery (InXight) which help the participant to navigate through complex knowledge structures and also handle the links from the leaves of the hyperbolic tree to the actual knowledge elements,
   - *integrated presentation of knowledge elements in knowledge maps*: knowledge maps are graphical representations of knowledge and its relation to organizational concepts. Examples are maps about knowledge holders, knowledge sources, knowledge structures, knowledge stocks, knowledge flows, knowledge processes, knowledge application or competence cards. They are used to visualize relationships between knowledge elements and their providers in an organization (e.g., Eppler 1997, Probst et al. 1998, 107ff, Vogt 1998). With this function, knowledge maps are integrated into the KMS and can be used in the search for knowledge elements,
   - *presentation of semantic closeness between knowledge elements (semantic net)*: a number of tools use advanced text mining technology to analyze documents and visualize the semantic relationships between the documents. An example is the tool ThemeScape (Cartia) that clusters documents with similar contents together with the distance between two documents visualizing their semantic closeness. Themes are represented as “mountains” of documents,
   - *presentation of access paths to knowledge elements/knowledge clusters*: the tool Answer Garden analyzes the paths (e.g., links in an Intranet’s web of hypertext documents) which participants use to access knowledge elements. These access paths are displayed as trails to knowledge elements. The more participants have used a certain access paths, the more pronounced is the visualization of the corresponding trail (“beaten tracks”),
4. **Presentation of knowledge elements after the search (=search results):**
   - *ranking of knowledge elements*: search results are presented in an order which reflects either how closely they match the participant’s search term or the “importance” of the knowledge element which might be calculated using criteria such as publication date, number of links to this knowledge element

(citation score), or criteria known from collaborative filtering, such as number of accesses by different groups of participants (e.g., experts), subjective evaluations by e.g., subject matter specialists or by participants with a similar profile and the like (e.g., Autonomy KM Toolsuite),

- *presentation of full texts*: search results regularly consist of a list of titles of matching knowledge elements, sometimes including the first paragraph of the description of the knowledge element (e.g., Hyperwave Information Server) or a short summary describing the knowledge element (e.g., Open Text Livelink). In case of documents this function allows that documents of varying formats can directly be viewed within the KMS (e.g., with so-called viewers),
- *presentation of related knowledge elements*: this is again a function of collaborative filtering. The KMS compares participants' profiles and suggests knowledge elements which the participant had not searched for, but which other participants who have a similar profile and who previously had got the same results had also searched for ("Who searches for X, searches also for Y"). An example for this function is Amazon.com's service that suggests a list of "related books" which builds on customers' profiles derived from their shopping history and other customer data,
- *navigation from knowledge elements to authors, experts or communities*: this function supports the interactive use of KMS. The participant can directly contact the author of a knowledge element, experts in the domain, subject matter specialists or knowledge brokers responsible for the corresponding topic or communities that discuss related issues e.g., by email or videoconferencing.

Finally, discovery services also comprise reporting about the state of the knowledge base, its use and users.

5. Reporting:

- *reports concerning knowledge elements*: can provide measures such as the number of accesses to each knowledge element, the number of searches with a certain term, the number of search results to specific search terms etc. In the case of no or a small number of search results to a popular search term, or a low ratio of organization-internal to organizational-external search results, these measures might suggest that knowledge in that area has to be developed[524],
- *reports concerning participants*: these functions monitor the patterns of usage of KMS by participants or collectives of participants. Examples for measures are the number of contributions in newsgroups, the number of knowledge elements published and the number of information subscriptions. These mea-

524. One interviewee reported that the monitoring of their KMS usage had revealed that the search term "Linux" had been searched more and more frequently by the participants. The organization decided on the basis of this information to offer Linux courses to a large number of employees.

sures can be used for motivation instruments such as a ranking of best knowledge providers in an organization or incentive systems, e.g., to reward the best contributors to every topic considered important for the organization. However, the design of incentive systems for knowledge management is a challenging task and experiences with simple measures such as the number of contributions to lessons learned data bases are not encouraging[525].

7.4.4 Publication services

This group represents *input-oriented functions of KMS*. Apart from a decentral publication of knowledge elements by the participants without support by technical staff, this group provides important functions for the organization of knowledge. Knowledge elements have to be linked with other knowledge elements as well as within the knowledge structure (ontology) of the organization. The quality of these functions has substantial influence on the quality of retrieved search results as knowledge elements that are not linked appropriately (i.e., corresponding to the collective mental models of the participants) are hard to find.

1. Knowledge publication:

- *publication of pre-structured contents by participants*: forms and templates provide guidance for the documentation of knowledge,
- *publication of not pre-structured contents by participants*: participants can store documents of all kinds of formats and structures in the KMS,
- *indexing/integration of published contents*: indexing helps the participant to provide a list of keywords to the published contents. The function integration of published contents is used to link a knowledge element to the organization's knowledge structure,
- *feedback from participants to authors of knowledge elements*: participants can provide structured or unstructured feedback to the author(s) of knowledge elements,
- *comments to knowledge elements*: participants can publish comments to knowledge elements ("post-it" function) which in turn can be used by other participants to cooperatively evaluate knowledge elements.
- *automatic notification of potentially interested*: once a new knowledge element is published, the KMS automatically selects and notifies participants who are potentially interested in the newly published knowledge. The selection of potentially interested participants might be based on information subscriptions, memberships in communities, organizational roles, an analysis of profiles or on the access history of participants,

525. The interviewee at Ernst & Young reported that his organization abandoned this practice years ago after their experience data bases were flooded with documents of questionable quality greatly reducing the signal to noise ratio. However, recently several organizations have started more sophisticated incentive programs for knowledge sharing, e.g., Siemens and Hoffmann La Roche.

2. **Knowledge organization:**
   - *development and management of knowledge maps*: knowledge maps are not developed separately from the KMS, but the KMS provides functions that help the knowledge manager to semi-automatically derive maps from the contents of the KMS. Examples are InXight Smart Discovery (InXight), SemioMap (Semio Corp.), ThemeScape (Cartia) and AnswerGarden2[526],
   - *knowledge repository*: a repository is a system used to store meta-data about objects of information systems such as data, functions, application systems, hardware, users or organizational units (Mertens et al. 1997, 345f). Knowledge repositories support the management of meta-information for knowledge elements (e.g., documents, authors, experts, communities),
   - *automatic indexing of full texts*: documents are scanned with text mining techniques that suggest a list of keywords for the texts which is compatible to the organization's knowledge structure (Grothe/Gentsch 2000, 212ff),
   - *automatic integration/classification/linking of knowledge elements*: again, text mining techniques are applied in order to e.g., discover interesting relationships between documents, classify documents, integrate them with the knowledge structure or cluster documents that cannot be integrated into the organization's knowledge structure. Thus, text mining provides techniques for a bottom-up document-driven categorization of knowledge elements which can be combined with a top-down categorization developed in e.g., an expert workshop (Grothe/Gentsch 2000, 217),
   - *semantic analysis of knowledge elements*: the KMS discovers relationships within and between knowledge elements. On the basis of techniques such as language analysis, semantic nets of terms are developed that describe a collection of knowledge elements,
   - *(hyper-)linking of published contents* (within documents): traditional documents (e.g., developed with text processing software such as MS Word) are transformed into hypertext documents in which hyperlinks are used to directly navigate within the documents, e.g., between sections of the documents or to cross-references,
   - *structuring and management of knowledge clusters*: the KMS provides functions to support the development and management of theme-specific knowledge areas or clusters containing knowledge elements to a specific topic.

7.4.5 Collaboration services

Apart from the advanced management of knowledge elements as described in the groups of services above[527], communication and cooperation is the second impor-

526. See also the function *integrated presentation of knowledge elements in knowledge maps* in section 7.4.3 - "Discovery services" on page 322.
527. See sections 7.4.3 - "Discovery services" on page 322 until 7.4.4 - "Publication services" on page 326.

tant part of a corporate KMS. Advanced support for organizational communication and cooperation regularly builds on a corporate Intranet and/or Groupware platform that supports basic functionality such as email or discussion lists[528]. Functions for knowledge communication and cooperation can be classified like general Groupware tools and systems according to time, location and flexibility of communication and cooperation (Koch/Zielke 1996, 70ff). In the following, time is used as criterion for the classification and distinguishes synchronous (all participants are on-line at the same time) from asynchronous communication and cooperation (simultaneous presence of communication partners is not necessary).

1. Asynchronous communication and cooperation:

- *email*: is the electronic pendant to traditional mail. Basically, ASCII text messages and so-called attachments (binary files) or MIME messages (Multi-Purpose Internet Mail Extensions) can be sent easily and quickly between email clients using the Internet or an organization's Intranet with the help of specific protocols, such as SMTP, the Simple Mail Transfer Protocol, POP3, the Post Office Protocol in version 3 or IMAP4, the Internet Message Access Protocol in version 4 (e.g., Höller et al. 1998, 10ff, Horn 1999, 42ff, Röckelein 1999, 40f),
- *email distribution lists*: are lists of email addresses maintained by the participants used to broadcast emails to multiple receivers.
- *listserver*: is a software tool that automates the management of mailing lists. The listserver can handle many mailing lists at the same time. Participants who want to join a mailing list simply send a message to the listserver (e.g.: subscribe list_name first_name last_name). The listserver then sends a request to the list's manager whether the participant should be added to the list. Every member of a list can send messages to the listserver which in turn are forwarded to all the members of the list either immediately or as a digested set of messages in defined time intervals, e.g., daily. Examples for listserver software are ListProc, Listserv (L-Soft) or Majordomo (Vaughan-Nichols1997, 162ff).
- *ad-hoc workflow management system*: workflow management systems primarily support well-structured organizational processes[529], but lack support for ad-hoc workflows (Koch/Zielke 1996, 158). Groupware platforms (e.g., Lotus Notes) and knowledge management systems (e.g., Open Text Livelink) offer this kind of flexible functionality.
- *newsgroups*: a newsgroup is a discussion list on a certain topic, a forum for exchanging ideas. NetNews is a public collection of newsgroups (more than 25,000) which are hierarchically organized according to themes (top-level themes are e.g., .comp for computer-related topics, .rec for recreational topics, .sci for scientific topics). Participants can subscribe to a selection of these

528. See also section 7.6 - "Classification" on page 361.
529. See section 4.3 - "Knowledge management systems" on page 82.

newsgroups[530] which might contain valuable information for certain groups of participants (e.g., system administrators, programmers). Additionally, organization-specific discussion lists can be set up with the help of tools which are part of Groupware platforms or Intranet solutions (Horn 1999, 46ff and 274ff, Grothe/Gentsch 2000, 78ff),

- *co-authoring functions*: support an asynchronous joint development of knowledge elements (e.g., documents) by multiple dislocated authors (Zwass 1992, 641). Examples for functions are version management, check-in, check-out of parts of (distributed) documents, highlighting updates attributed to a certain author, management of comments, accept/deny proposals for changes and the like. Examples for tools that provide co-authoring functions are document management systems such as Documentum 4i (Documentum) or Panagon (FileNET),
- *administration of group profiles and privileges*: this functionality supports in analogy to the definition of roles for participants and profiling for participants the definition of privileges and profiles for collectives of participants, such as work groups, teams and communities. The functions for administration of collectives of participants greatly support interaction within groups, collaboration, e.g., collaborative filtering and the transactive memory as groups can be connected to information flows in the same way as described for individual participants above. These advanced administration functions together with intelligent agent technology that uses participants' individual profiles as well as group profiles can provide the basis for an *intelligent community portal* (Grothe/Gentsch 2000, 267ff).

2. **Synchronous communication and cooperation:**
   - *point-to-point video conference*: also sometimes called "simple video conferencing" connects two participants and transmits motion pictures usually captured by a small video camera which is mounted on the participants' monitors as well as audio recorded by microphones between them via the Internet, ISDN (Integrated Services Digital Network) or the analogous telephone network, sometimes also referred to as POTS (Plain Old Telephone System, Horn 1999, 18ff, 227ff). Examples for video conferencing software are CU-SeeMe (White Pine) or NetMeeting (Microsoft) which offer a lot of additional functionality, such as whiteboard, application sharing, text chat and multi-point video conferencing,
   - *multi-point video conference*: in addition to point-to-point video conferencing a multi-point video conference involves more than two participants and thus requires a multicast capable network infrastructure (see Wittmann/Zitterbart 1999) or a multi-point control unit or service (Horn 1999, 231). Examples are the Polycom video conferencing tools (Polycom).

530. The organization can preselect those newsgroups it wishes to offer to its employees.

- *networked group video conferencing rooms*: in the beginning of tele-conferencing, many multinational organizations (e.g., IBM, Siemens) or telecommunication companies (e.g., Deutsche Telekom) installed the expensive video conferencing equipment in a separate room, a video conferencing studio (Hansen 2001, 431). The immobile installation of professional video equipment provides high quality pictures and sound. With the advent of cheap and powerful desktop video conferencing systems, the use of video conferencing studios is limited, e.g., for electronic group meetings of managers or of two and more geographically dispersed work groups or project teams,
- *audio conference*: is the electronic equivalent to the telephone. Two or more participants communicate via electronic networks. If the Internet is used as the communication medium, audio conferencing is also sometimes called Internet-telephony (Vaughan-Nichols 1997, 204ff, Horn 1999, 223ff). Examples for audio conferencing tools are Surf&Call (VocalTec) or the audio part of NetMeeting (Microsoft),
- *group conference management*: functions for a management of tele-conferences support a person to moderate a group discussion. A dedicated moderator can for example restrict access to certain participants, ban unwanted contributions or participants, thread contributions, administer votings and the like. Examples for software tools supporting the moderation of text chats are the moderation module of SpinChat (Spin) or the MBone moderation tools (Malpani/Rowe 1997, Perry 1997, 13ff),
- *instant messaging*: is the synchronous form of email. A participant can send a text message to a person (or a group of persons) that is delivered immediately. The best known systems supporting this function are ICQ–"I seek you"[531], the AOL Instant Messenger[532] (see Horn 1999, 49) or the Microsoft Messenger[533],
- *chat*: is a form of text-based tele-conferencing. A chat is a synchronous forum for discussions which displays all contributions immediately after they have been typed by the participants. Internet-based public chat server, so-called IRC (Internet Relay Chat), offer thousands of theme-specific channels, so-called conference rooms, and are visited by many thousand people daily. Web portals such as YAHOO! (URL: http://www.yahoo.com/) or web.de (URL: http://chat.web.de/) offer overviews over chat offerings (Horn 1999, 48ff). Apart from these public on-line discussion groups many organizations internally use chat software to support text-based conferences about certain topics or as a brainstorming tool. One example for commercial chat software applied in organizations is SpinChat (Spin).
- *electronic whiteboard*: is part of a tele-conferencing system. It offers functionality similar to a simple paint software (e.g., Windows Paint) that can be

531. See URL: http://www.icq.com/
532. See URL: http://www.aol.com/
533. See URL: http://messenger.msn.com/

used simultaneously by multiple dislocated participants of a tele-conference to share information, import and jointly work on documents, drawings or images and the like (e.g., the whiteboard in Microsoft NetMeeting, Horn 1999, 233),

- *application sharing*: is a form of tele-conferencing where several dislocated participants jointly use an application and simultaneously work on e.g., CAD designs, spreadsheets, graphs or text documents (Hansen 2001, 431f). One popular example is the application sharing functionality offered by Microsoft NetMeeting in connection with Microsoft's Office applications (e.g., Excel, Powerpoint, Word and Access, Horn 1999, 233f),
- *electronic brainstorming*: is a specific function that is often part of Groupware tools. Brainstorming tools usually support generation and organization of ideas. The software enables participants to submit ideas to a topic and immediately presents these ideas to other participants. One example for brainstorming software is GroupSystems (Valacich et al. 1991),
- *list of participants currently on-line*: are also called "Buddy lists" and an instrument to increase awareness of what is going on in a KMS. Tele-conferencing, no matter whether text-based, audio or video conferencing, requires that participants are on-line. In order to support the initiation of tele-conferences, participants need to know who else in an organization is on-line and on which computer they are. Due to data privacy laws, access to a list of participants currently on-line is regularly restricted. Additionally, in large organizations the participant might need additional information about the other participants in addition to a person's login (e.g., name, location, position, roles, memberships, competencies etc.) and navigation help (e.g., find all participants on-line worldwide who work for a specific business process),

7.4.6 Learning services

As mentioned before, the market for KMS in general develops from advanced document management systems and thus a focus on explicit, codified knowledge to the integration of collaboration and e-learning functionality and thus a focus on implicit, personalized knowledge (see also the empirical results in part C). E-learning suites provide a basis for an organization-wide integrated management of CBT and WBT modules and also for computer-supported cooperative learning (CSCL) or distributed collaborative learning (Möhrle 1996). Examples for elements of an e-learning suite such as Lotus's LearningSpace are: administration of course materials and (hyper-) media, e.g., documents, audio and video files, links etc.; a schedule that provides an overview of programs, courses, times etc.; a so-called course room for on-line exchange of ideas and discussions between students and teachers; profiles of participants and administration of exams (e.g., Lehner 2000, 389f, Seifried/Eppler 2000, 33).

1. Asynchronous CBT and tele-learning:

- *computer based training*: this function supports the integrated and context-dependent access to CBT modules within KMS. Examples for software that specifically focuses on tele-teaching and tele-learning in organization's Intranets are so-called e-learning suites such as LearningSpace (Lotus) and the Hyperwave E-Learning suite (Hyperwave),

- *video server*: is in analogy to a data base server a computer system that stores and handles accesses to video files. Video server have been heavily discussed in connection with *video-on-demand* (e.g., Röckelein 1997, 56f, Hansen 2001, 114). Video server provide functionality so that participants can access any video file (e.g., a lecture, a product presentation, a penal discussion) at any time. Video streaming server allow that the user does not have to wait until the entire file is loaded, but can already watch the video while the file is loaded,

2. Synchronous CBT and tele-learning:

- *live broadcasting of videos*: this functionality is the synchronous equivalent of video servers. It supports the broadcasting of e.g., lectures, presentations or the CEO's weekly talks to participants who have to be on-line and "tuned in" at the time of the broadcasting. Video broadcasting is applied in tele-teaching (e.g., at the two-campus University of Erlangen-Nürnberg) and in business TV (Lehner 2000a, Weidler 2000). Recent implementations regularly include functionality to support interaction between the receivers of the broadcast and the sender (feedback channels for text, audio or video). Examples for software tools supporting video broadcast are the MBone tools and IntraTV (Siemens Business Services, Lehner 2000a, 15f).

Generally, it is supposed that organizations with a KMS solution (no matter whether bought on the market or developed internally) have implemented a larger number of KMS functions than organizations without a dedicated KMS solution. This should be especially true for the more advanced KMS functions which are not available as part of a basic Intranet or Groupware platform. Consequently, the following hypothesis will be tested:

Hypothesis 18: Organizations with KMS have a larger number of KMS functions than organizations without KMS

KMS architectures also strongly aim at an integration of existing data and knowledge sources as well as existing knowledge-related services (e.g., documentation, visualization, search and retrieval as well as collaboration). Thus, there should also be a positive correlation between the existence of a KMS in an organization and the integration of KMS functions. This should be especially true for KMS bought on the market, because according to interviews with vendors of KM suites as well as knowledge managers applying such systems, integration across platforms and formats is the single most important reason why organizations invest in KMS available on the market. This leads to the following hypothesis:

Hypothesis 19: KMS functions in organizations with KMS bought on the market are more integrated than KMS functions in organizations without KMS

7.4.7 Personalization services

Subject matter specialists and knowledge brokers are responsible for e.g., the refinement of knowledge, the distribution of knowledge to potentially interested members, for the identification of trends in the use of KMS, for the acquisition of external knowledge about topics that are needed or for the motivation of participants to contribute[534]. The functionality within the group administration of KMS supports these specific roles in their tasks, but also individual participants and groups, teams and communities in the personalization of interfaces and knowledge managers in monitoring the system usage as measures of success[535].

Knowledge push functions can generally be initiated either decentrally by the participants or centrally e.g., by subject matter specialists or knowledge brokers who can therefore easily distribute information to interested groups of participants:

- *profiling*: participant profiles contain general information about a participant such as job description, roles, privileges, interest profiles or the level of experience which are used to narrow the search domain and improve the relevance of search results. Consequently, KMS have to extensively apply complex user models in order to provide this kind of support (Mertens et al. 1997, 53f, Mertens/Höhl 1999). Profiles can either be administered by the participants themselves or centrally by knowledge managers, subject matter specialists or knowledge brokers,
- *information subscriptions for interested users*: the participant subscribes to an information service which will automatically send personalized messages in certain time intervals or event-triggered. The messages contain information and/or links to information that match the participant's profile. A recent development in the field of information subscriptions are so-called *news channels* or *news ticker* which permanently display news in a separate line e.g., at the bottom end of the screen (for examples for news ticker and information subscriptions on the Internet see Horn 1999, 62ff),
- *intelligent (search) agents*: the term *agent* in general denotes an autonomous piece of software that carries out actions for a user (Mertens et al. 1997, 6). Technologically, agents are based on approaches of distributed artificial intelligence. Like information subscriptions intelligent search agents automatically search in knowledge bases for information that matches a predefined participant's profile. Additionally, agents can e.g., negotiate with other agents in other systems to provide more intelligent search results and learn about the participant to extend his or her profile according to the history of searches and evaluation of search results,
- *personalization of user interface*: in order to avoid "information overload" due to the abundance of organization-wide knowledge resources, many KMS offer functions to personalize the participants' interface with the system,

534. See section 6.1.2 - "Knowledge management roles" on page 162.
535. See also chapter 8 - "Economics" on page 395.

sometimes called *my place* (e.g., Grothe/Gentsch 2000, 73f). The idea of a personalized, individual window to the organization's knowledge assets and applications is closely connected with enterprise information portals. *Enterprise information portals* (EIP) or corporate portals offer e.g., enterprise-wide search functionality, navigation, directory browsing as well as links to external Web sites and information sources (e.g., Kappe 2000). EIP software offers functionality for personalization so that every participant accesses the KMS with the help of an individual information portal. Examples for EIP software are the Hyperwave Information Portal (Hyperwave) or the E-Portal (Viador),

- *definition of roles for participants*: in analogy to networks and data base management systems, KMS administration can be greatly supported by the concept of roles. One individual employee can play several roles with respect to the use of KMS, e.g., various *functional roles*, such as consultant, sales person, engineer, member of R&D, *knowledge-related roles*, such as subject matter specialist, knowledge broker, knowledge manager as well as the *role of a technical administrator* of the KMS etc. On the one hand, roles determine the participants' privileges, e.g., for accessing, publishing, updating and deleting knowledge elements etc. On the other hand, roles can be used to narrow the search domain and help to navigate the organization-wide knowledge structure,
- *role-specific configurations of knowledge management systems*: roles can further be used as the basis for a pre-configuration of KMS. Specific groups of participants get predefined default parameters, e.g., for the selection of topical data bases in a Lotus Notes environment, specific on-line help or role-specific lists of experts, networks and communities. Examples are special configurations for newly recruited versus senior management bundles. Trainees at Andersen Consulting for example get a so-called starter package dependent on the trainee's educational background (e.g., IT versus business). The starter package pre-selects a number of data bases, news feeds, membership in communities etc. potentially interesting for the trainee, arranges them on the participant's screen and provides special instructions for the use of the KMS.

7.4.8 Access services

The KMS services described so far are accessed by a variety of access services. The simplest way to access a KMS is via a standard Web browser (e.g., Microsoft Internet Explorer, Mozilla, Netscape Navigator[536]). However, more advanced KMS have lived up to the requirement that KMS have to be seamlessly integrated into the ICT work environment that the knowledge worker chooses to use. This integration requires the following groups of functions:

- *transformation and translation to other applications*: access services translate and transform the contents and communication to and from the lower levels of

536. URLs: http://www.microsoft.com/, http://www.mozilla.org/, http://www.netscape.com/

services in the KMS to heterogeneous applications. Examples for applications are a Web browser, a file management system (e.g., Microsoft Windows Explorer), an email client (e.g., Eudora Email, Netscape Mail[537]), personal information management applications such as calendar, to-do lists, address books (e.g., Microsoft Outlook, Palm Desktop[538]) as well as collaboration or Groupware platforms (e.g., Lotus Notes/Domino, Microsoft Exchange[539]).

- *transformation and translation in mobile environments*: knowledge workers have often advanced demands for mobility, thus access services also have to cope with varying communication environments, especially bandwidths anywhere between a fast local area network and rather slow telephone lines or even offline work with KMS contents and therefore replication of (parts of the) contents (see integration services). Examples for appliances that are used to access KMS are a PC, a notebook, a personal digital assistant (PDA), a smartphone or a feature phone. These appliances differ with respect to their resources, e.g., screen size, processing power, storage, or interaction capabilities. Contents have to be transformed, so that they can be handled with the different appliances.
- *integrated knowledge workspace*: synchronization between these different applications and appliances including an integrated management of meta-data can be provided by an integrated knowledge workspace (see also integration services). In its simplest form, the knowledge workspace can be thought of as a portal that provides access to the most important applications that the knowledge worker works with (see also personalization services). A more advanced knowledge workspace would be aware of the knowledge objects that a knowledge worker accesses in different applications and relate them to each other on the basis of an extended meta-data management (meta-data brokering, ontology brokering).
- *authentication and authorization*: the KMS have to be protected against eavesdropping and unauthorized use by tools for authentication, authorization and encryption. KMS in many cases have to be accessed not only from within the boundaries of a corporate LAN, but also from outside via telephone lines, e.g., using a remote access system, and/or the Internet which requires strong encryption. In this case, access and infrastructure services together have to be in place to provide secure access to the corporate KMS. An example is the use of a virtual private network (VPN) that realizes a kind of a secure "tunnel" through which data are transferred to and from the corporate KMS, e.g., using the point-to-point tunneling protocol (PPTP). Access has to be restricted to those knowledge objects that the knowledge worker is allowed e.g., to see[540],

537. URLs: http://www.eudora.com/, http://www.netscape.com/
538. URLs: http://www.microsoft.com/, http://www.palm.com/
539. URLs: http://www.lotus.com/, http://www.microsoft.com/
540. "See" in this case means that the existence of a knowledge object is made known to the knowledge worker, but she is not allowed to access the contents of the knowledge object, e.g., a hypertext document.

to view its contents, to download, to change, to add versions to, to delete etc. (see also infrastructure services).

7.4.9 Example: Open Text Livelink

Open Text's product family Livelink represents one of the leading KMS platforms with a centralized architecture. Livelink has a large installed base of millions of users in 46,000 organizations in 114 countries[541] many of which are large organizations. Figure B-60 assigns Open Text Livelink's modules to the six layers of the centralized KMS architecture (see Figure B-59 on page 319). In the following, Livelink's components are briefly discussed.

Data and knowledge sources. Livelink data is stored in a relational data base system and the file system of the server's operating system on which Livelink is installed. Various other data and knowledge sources are made available by services on the infrastructure layer covering structured as well as semi-structured, organization-internal as well as -external sources.

Infrastructure services. Livelink is based on the organizational Intranet infrastructure. On the infrastructure level, it offers functionality for administration, workflow as well as import and export of XML data. Open Text offers a large number of modules targeted at enhancing technical access to the system (WebDAV, eLink, Directory Services, Remote Cache), integration with other Livelink instances or Open Text products (Brokered Search, Doorways, Collections Server Integration, Library Management Integration, GISLink, DocuLink) and external systems (Spider), easing administration of the Livelink server (Performance Analyzer, Monitoring Agent, DB Backup Validator, Object Importer, Recycle Bin) as well as enabling or supporting development of individual extensions based on the system's API (SDK, XML Workflow Interchange/Extensions).

Integration services. In Livelink, knowledge is stored in and represented by so-called "objects", e.g., documents, folders, discussions, news channels or task lists. All of them can be placed in a folder hierarchy that resembles traditional file systems. Meta-data is either added automatically, e.g., creation/change date, creator, protocol, or manually via customizable categories. Because all meta-data are stored in a relational data base, it can be queried using SQL statements in so-called reports. Optional modules offer functionality for manual or automatic creation of multiple alternative taxonomies (Classifications Professional, Taxonomy Workbench), extensions of the meta-data model (Attribute Extensions) and securing user information (Privacy Panel).

---

541. According to Open Text investor relations; see also: URL: http://www.opentext.com/investor/. With these figures, Livelink claims to be the largest independent provider of what it calls now enterprise content management solutions.

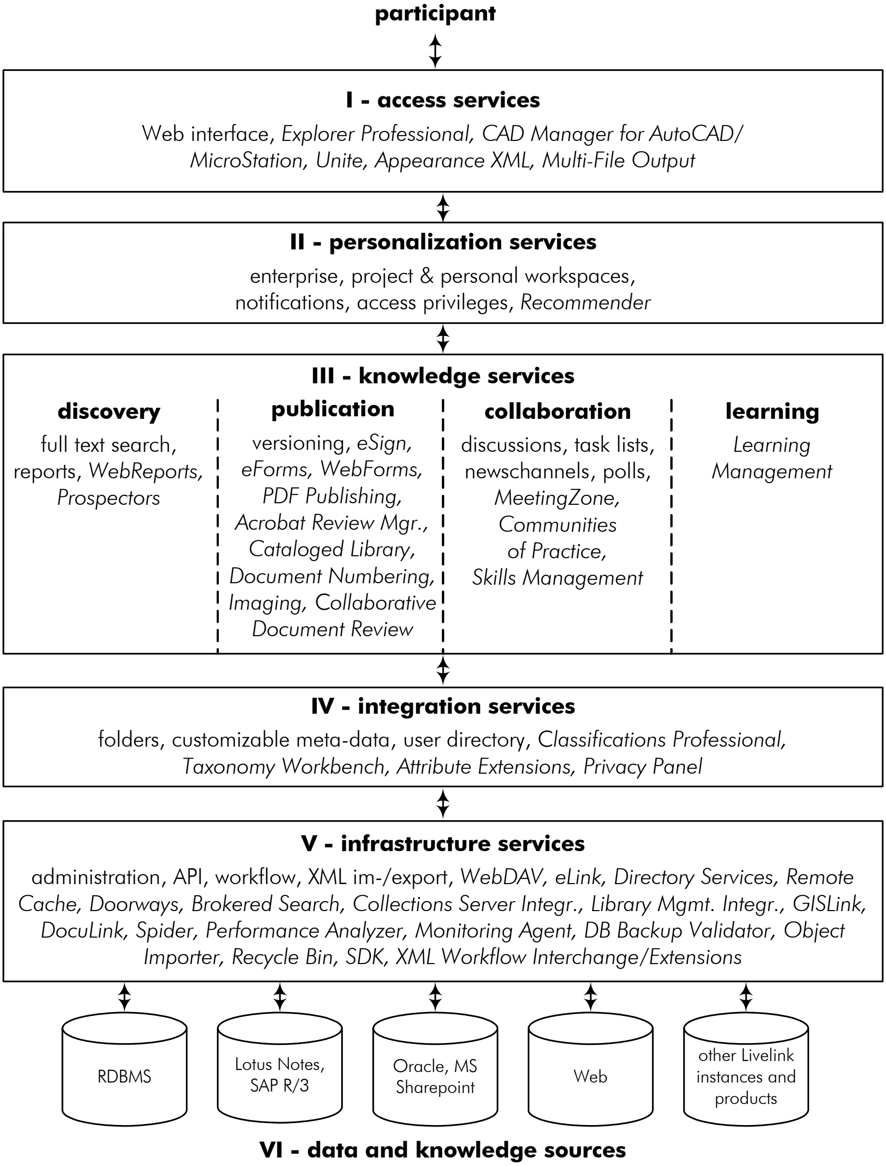

FIGURE B-60. Open Text Livelink in centralized KMS architecture[542]

542. Italic descriptions refer to separate software modules that extend Livelink's core functionality. It depends on the actual license agreement whether they are included or not. A variety of additional modules can be obtained from 3rd party vendors and are not considered here.

Discovery services. Livelink's full-text search engine allows basic and advanced keyword searches. Additionally, assigned meta-data can be used for limiting the search domain. A typical search result page not only includes a ranked list of various types of objects with short descriptions, e.g., documents, discussion topics, folders or objects from further knowledge sources made accessible through additional services on the infrastructure level, but also gives hints to what authors or creators have been most active according to the actual query. Livelink's notification mechanism allows users to place change agents on selected folders to be notified via email if changes occur, e.g., if a new document or a new version of a document is added. Comparable functionality is provided by an optional module to monitor changes of search queries to Web resources (Prospectors). Another module is available for generating reports on structured data, e.g., forms, external data bases or Livelink management data (WebReports).

Collaboration services. Some basic functions like discussions (black boards), polls, news channels, task lists, and advanced functions like workflows aim at supporting collaboration. MeetingZone comprises a set of meeting support tools that are integrated into Livelink, e.g., a whiteboard, a chat, a shared desktop and objects to be used during the meeting. Other optional modules offer basic functions for competence management (Skills Management) and support creation of community workspaces (Communities of Practice).

Learning services. The standard version of the Livelink server does not provide any learning services. Structured course units as well as question and answer tests can be created by means of an optional module that also allows for integration with the Skills Management module (Learning Management).

Publication services. Typical document management functions of Livelink are check-in/check-out of documents in order to avoid conflicts if more than one user works with a document, a versioning mechanism and workflows that support publication of documents, e.g., a release workflow. All types of files can be stored in Livelink. The most common types, e.g., formats of office systems, can be converted to HTML on demand. Thus, documents can be viewed without the native application and indexed by Livelink's search engine. Optional modules provide capabilities for electronic signatures (eSign), management of electronic forms (eForms, WebForms), creation of portable document format (pdf) files from within Livelink (PDF publishing) and managing linked annotations (Review Manager for Acrobat), administration of bibliographic resources (Catalogued Library), labelling documents (Document Numbering), and coordinating the steps of the document review processes (Collaborative Document Review).

Personalization services. Livelink offers three types of workspaces that differ mainly with respect to what groups of users are granted privileges to access them. The enterprise workspace is the central workspace for all users. A personal workspace belongs to every user with access restricted to this user. Project workspaces can only be accessed by participants defined by the project's coordinator(s). The operations users and groups of users may perform on an object are defined by detailed privileges at the granularity of single objects. Examples for operations are: see object, see content of object, delete object, change meta-data or add version. All knowledge services, e.g., discovery services, as well as access services consider these privileges. An additional module generates suggestions of potentially interesting contents based on the individual user profile (Recommender).

Access services. The standard way to access the system is with the help of a standard Web browser, e.g., Microsoft Internet Explorer, Mozilla or Netscape Navigator.[543] Thus, access to Livelink is relatively platform-independent and not limited to a corporate LAN. Due to the fact that access to Livelink requires no additional installations, e.g., of plug-ins[544], Livelink can also be accessed via the Internet from every networked computer that has a Web browser installed. Nearly all objects stored in Livelink can be exported to and imported from XML documents. Additional modules integrate Livelink with emails (eLink) or desktop applications such as MS Windows Explorer, MS Word, Adobe Acrobat, and CAD applications (Explorer Professional, CAD Manager for AutoCAD/MicroStation). The Explorer provides a drag & drop integration into the Microsoft Windows Explorer. A "professional version" extends this integration with basic online/offline synchronization functions and an integration e.g., into Microsoft Office. An example is check-in/check-out of documents directly from Microsoft Word. Other optional modules provide access to and consolidate multiple Livelink instances on the presentation layer (Unite), allow for adaptations of the Web interface (XML Appearance) and for printing, mailing or downloading multiple files over the Web interface (Multi-File Output).

Figure B-60 categorizes the most important functions and modules (in italics) of Open Text's core product, the Livelink Enterprise Content Management (ECM) Enterprise Server 9.7.0[545]. After acquisition of several companies such as Artesia, Gauss, IXOS, RedDot and, more recently, Hummingbird, the software provider offers an even larger variety of different products and variations of the Livelink ECM server under the umbrella of the Livelink ECM family and addresses topics such as

- KM and collaboration: Livelink ECM Knowledge Management, Collections Server, Discovery Server, Federated Query Server, Library Management, Collaboration,
- Web content management: RedDot Web Content Management, Livelink ECM Web Content Management Server,

543. URLs: http://www.microsoft.com/, http://www.mozilla.org/, http://www.netscape.com/
544. However, the comfortable use of Livelink requires the installation of Java Virtual Machine.
545. Available since December 2006.

- email archiving and management: Livelink ECM Email Archiving / Monitoring for Lotus Notes, Email Archiving / Monitoring / Management for MS Exchange),
- compliance, governance and archiving: Livelink ECM Internal Controls, Records Management, Regulated Documents, Accreditations Server, Collaborative Submissions, Litigation Management, Content Lifecycle Management, Contract Lifecycle Management, Document Management, Library Management, Archiving for File Systems, MS Sharepoint Integration,
- high volume document processing and imaging: Livelink ECM Production Document Management, Production Imaging,
- digital asset management and publishing: Artesia Digital Asset Management, Livelink ECM Enterprise Publishing,
- content and document management in public institutions: DOMEA Government Content Management product family,
- document management and archiving with SAP: Livelink ECM Suite for SAP solutions,
- connectivity between software platforms: Hummingbird Exceed product family, Hummingbird Security, NFS Maestro, HostExplorer product family,
- extensions for Oracle: Livelink ECM Accounts Payable for PeopleSoft Enterprise, for JD Edwards EnterpriseOne, and for JD Edwards World,
- business process management: Livelink ECM Advanced Workflow, Business Process Management Server,
- project management: Livelink ECM Clinicals, Construction Management, Program Management,
- reports: Livelink ECM Report Output Management, Vista Plus Suite,
- technology-enhanced learning: Livelink ECM Eloquent Media Server,
- portal integration: Livelink ECM Portal Integration Kit.

A recent addition to the product portfolio is Livelink ECM eDOCs, formerly the Hummingbird Enterprise Suite, a complimentary product family for document management, records management, contract management, correspondence management which also offers functions for collaboration, search, and workflow management and can be integrated with other products such as MS Sharepoint or MS Office[546].

Summing up, Open Text Livelink can be considered as a knowledge management system in the sense of a platform that combines and integrates a substantial number of functions for every level distinguished in the centralized KMS architecture. With roots in document management, Open Text Livelink's focus is on explicit knowledge, its publication and discovery across formats, platforms and the boundaries of a corporate LAN. Moreover, Livelink supports collaboration based

546. see http://www.opentext.com/, esp. http://www.opentext.com/2/sol-products.htm.

on co-authoring and sharing of documents. Livelink implementations can be found in many large organizations in Europe and the US. Although Livelink can be used (almost) out-of-the-box as a basic KM platform, most implementations adapt the user interface to corporate style guides and extend the integration and infrastructure capabilities of Livelink to cover organization-specific data and knowledge sources.

7.5 Distributed architecture

For quite some time, the only architecture that was discussed for KMS was a centralized one. This is due to the fact that a primary challenge for organizations has been to collect, organize and provide access to the pool of documented knowledge that is spread across a multitude of data and knowledge sources stored on a number of heterogeneous server systems and even on file systems of individual PCs. Centralized KMS provide a powerful instrument to consolidate the often fragmented organizational knowledge base. However, centralized KMS solutions require powerful machines, optimized software, i.e. a standard KM suite or an individual KMS software, and a lot of effort to tap into the multitude of existing data and knowledge sources and to semantically integrate them. Therefore, establishing a KMS with a centralized architecture is a costly approach.

Recently, the peer-to-peer metaphor has been discussed intensively as an alternative to server-based solutions that makes better use of the often abundant idle computing and storage resources that can be found in many organizations due to the fact that PCs have become powerful machines that provide abundant unused capacities.

In the following, section 7.5.1 reviews the peer-to-peer metaphor and section 7.5.2 discusses its application to KMS. Finally, section 7.5.3 presents Infotop, a peer-to-peer KMS that also targets another unresolved question in the design of KMS, namely the integration of KMS functions into the knowledge worker's personal workspace management.

7.5.1 Peer-to-peer metaphor

The term *peer-to-peer* denotes the idea of a network of equals (peers) that provide resources such as CPU time, storage area, bandwidth or information to each other so that collaborative processes are enabled avoiding a central coordinating instance (Schoder/Fischbach 2002, 587, Schoder et al. 2002). Ideally, peer-to-peer networks can be described by the following characteristics (Barkai 2001, 4ff, Schoder/Fischbach 2002, 587):

- *mutual client-server-functionality*: each peer can act as a client and as a server, thus rendering all nodes functionally equal,
- *direct exchange between peers*: there is no central node which coordinates the communication between the peers,
- *autonomy*: the peers are solely responsible for their activities, especially for determining what resources they share when and with whom.

In the terms of the client-server architecture, each peer, i.e., each computer participating in a peer-to-peer network, can act both as a client and as a server in the context of some application (Barkai 2001, 4). The peer-to-peer idea is not new, some argue that it is one of the oldest architectures in the ICT and telecommunication domain with the telephone system, the Usenet and the early Internet as major examples that employ this metaphor (Schoder/Fischbach 2002, 588). However, it is only recently that the peer-to-peer metaphor has received a lot of attention. The metaphor has been adopted in various application areas. Examples for application areas of existing peer-to-peer networks are[547]:

1. *instant messaging*, e.g., the well-known ICQ[548] network,
2. *file sharing*, with prominent examples, e.g., Gnutella, Kazaa, Napster or Overnet (Edonkey 2000)[549], i.e. peer-to-peer software that supports the sharing of files in networks of users, especially audio and video data as well as computer games,
3. *distributed and grid computing* which aims at a coordinated usage of distributed computing power, with the prominent example of the world-wide network that jointly processes data on the search for extraterrestrial life (SETI@HOME[550]),
4. *collaboration and Groupware*, with Groove[551] being the most cited distributed Groupware platform that employs the peer-to-peer-metaphor.

In the following, based on the ideas and developments in the fourth application area, collaboration and Groupware, the peer-to-peer-metaphor is applied to the complex area of knowledge management systems, called distributed or peer-to-peer KMS.

7.5.2 Peer-to-peer knowledge management systems

Recently, there are several attempts of KM researchers to profit from the promised benefits of a peer-to-peer metaphor in the design of an information sharing system and especially of a knowledge management system[552].

The following two figures, Figure B-61 and Figure B-62, together illustrate an ideal architecture of a peer-to-peer KMS or an extended peer-to-peer KMS respectively[553]. Figure B-61 shows a number of peers that together form a distributed knowledge management system. The peers are physically connected, e.g., via local area networks, telephone lines or the Internet. The connections are visualized by

547. Examples for existing application software realizing peer-to-peer networks more or less deviate from the ideal architecture. Most networks are supported by servers, or super peers, that aid e.g., awareness or localization of peers.
548. See URL: http://www.icq.com/
549. URLs: http://www.edonkey2000.com/, http://www.gnutella.com/, http://www.kazaa.com/, http://www.napster.com/
550. This project was initiated by the Space Sciences Laboratory of the University of California, Berkeley; URL: http://setiathome.ssl.berkeley.edu/
551. URL: http://www.groove.net/
552. Examples are Parameswaran et al. 2001, Benger 2003, Susarla et al. 2003, Maier/Sametinger 2004.
553. See also Maier/Sametinger 2004, Maier/Hädrich 2006.

solid gray lines. The dashed black lines show some examples of knowledge work processes that are supported by the peer-to-peer KMS. They include Nonaka's four processes of knowledge conversion, *externalization*, *internalization*, *combination* and *socialization*.

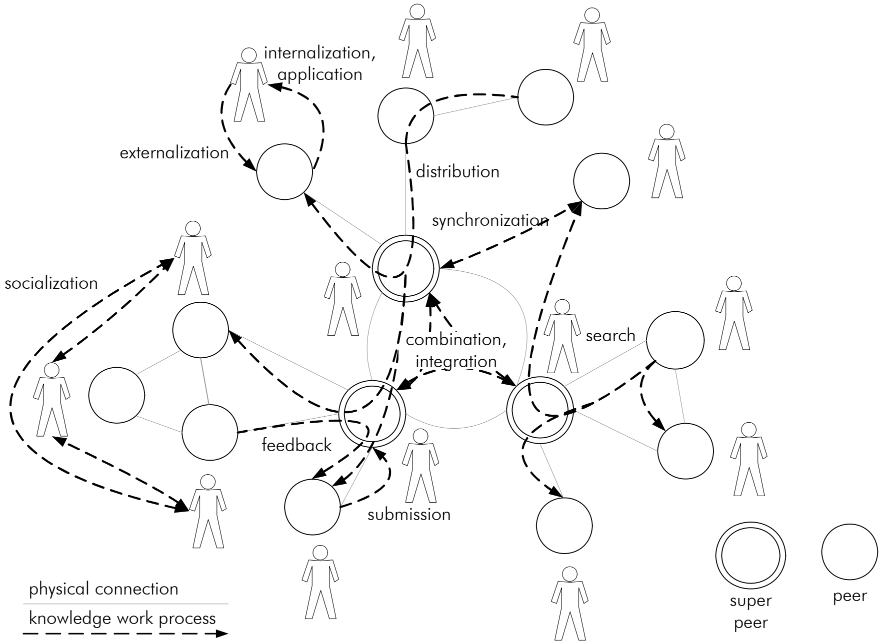

FIGURE B-61. Architecture of a peer-to-peer KMS[554]

In the following, the knowledge work processes shown in Figure B-61 are described shortly:

- *socialization*: as shown in the figure, socialization is only marginally supported by KMS, e.g., maintaining social relations over distances by instant messaging.
- *externalization*: knowledge is externalized, i.e. documented, contextualized and stored as explicit knowledge in one peer's individual knowledge base. This peer can now decide what other (groups of) peers should have access to this knowledge element.
- *internalization/application*: the reverse process also only involves the individual's personal knowledge base in its simplest form.
- *combination:* knowledge from several peers can be brought together semi-automatically or manually and stored as part of one, many or all the knowledge bases of the peers involved in the combination.

554. See also Maier/Sametinger 2004, Maier et al. 2005, 367.

- *distribution*: the distribution process means that knowledge is pushed from one peer to a certain group of other peers who can decide whether to accept the offered knowledge element(s) into their own knowledge bases.
- *search*: a search process can involve an individual's personal knowledge base as well as all the portions of other peers' knowledge bases (1) to which access has been granted and (2) which are accessible at the period of time when the search is performed.
- *feedback*: individuals can get feedback on their knowledge from any other peer who was granted access to that knowledge.

In many organizations, specific KM roles, such as a subject matter specialist or a knowledge (base) administrator, are established in order to e.g., collect, review, value, organize, store, refine or distribute knowledge that can then be reused by knowledge workers. In the ideal architecture of a peer-to-peer KMS shown in Figure B-61, these special roles are visualized by so-called "*super peers*".

Generally, super peers provide the same functionality as peers do. Every peer may act as a super peer and provide services of a subject matter specialist for a certain (set of) topic(s). The differences are that super peers also provide quality management to the distributed KMS architecture, improve performance of the network, increase accessibility of the workspaces and aid collaboration between the peers. Thus, super peers might provide a (large!) knowledge base that acts as a "knowledge cache" for a certain network segment. This reduces network traffic when peers from the same network segment repeatedly access certain knowledge elements from other peers in other segments. Specifically, super peers might provide the following services:

- *synchronization*: peers that sometimes work offline might subscribe to synchronization services offered by a super peer and thus improve their share in a peer-to-peer KMS and at the same time improve their network visibility even though they might be sometimes unavailable.
- *submission*: also, a submission process might be institutionalized by which every peer can push knowledge towards a subject matter specialist or knowledge base administrator respectively in order to get it reviewed, commented and, if accepted, get its quality certified. Possibly, meta-data on the knowledge element is also organized as part of the collection of (links to) knowledge elements that is administrated by the subject matter specialist.
- *integration*: super peers might also establish a joint effort to provide a standardized taxonomy or ontology of the knowledge domains that they are involved in and thus contribute to the integration of the diverse knowledge bases connected in the distributed KMS architecture.

Consequently, super peers ideally are powerful machines with abundant resources, a fast connection to the network and always online. Figure B-62 shows the architecture of a peer and a super peer in detail.

Both architectures basically consist of the same layers as the architecture of centralized knowledge management systems, but lack a centralized knowledge structure, taxonomy and repository. Thus, in the following only the differences to the centralized architecture are discussed[555].

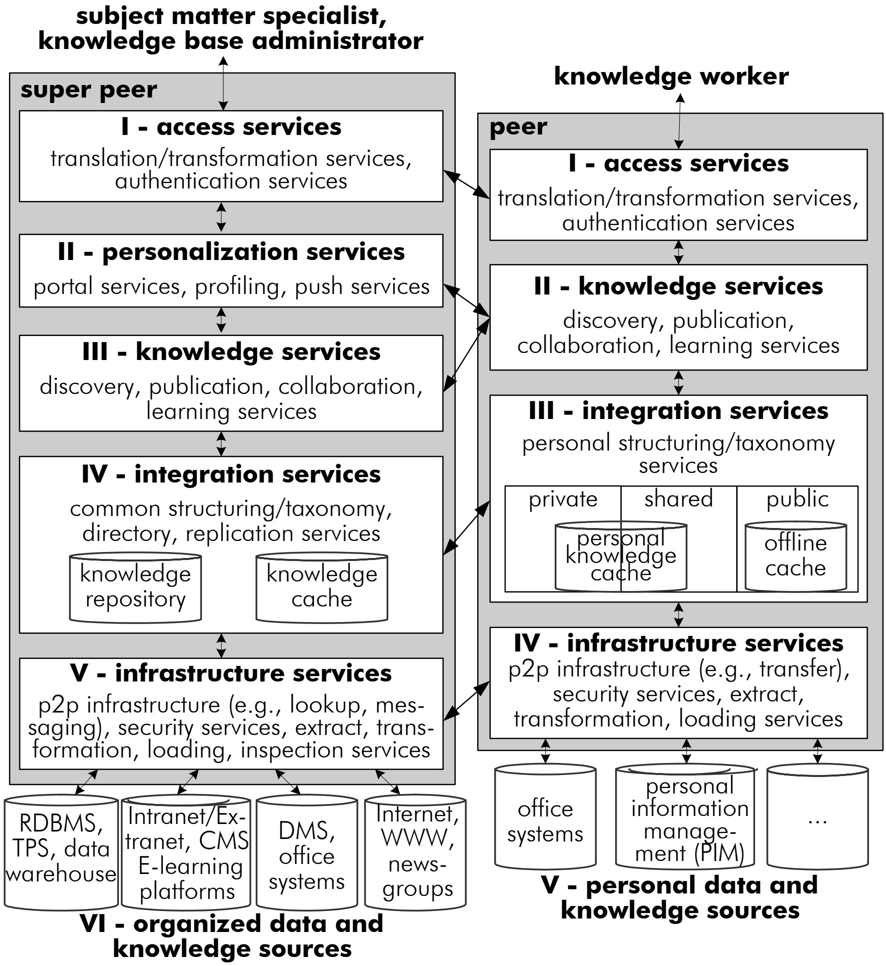

FIGURE B-62. Architecture of peer and super peer

Peer. The peer's architecture builds on infrastructure services that basically handle (1) extract, transformation and loading from personal data and knowledge sources and (2) provide the peer-to-peer infrastructure for locating peers, exchanging data with other peers and assuring security of the personal knowledge base. Integration services handle meta-data of the knowledge objects in the personal knowledge base

555. For the centralized architecture see section 7.3.3 - "Integrating architectures for KMS" on page 311.

and establish a personal knowledge structure or taxonomy. The knowledge base comprises private, protected and public areas. Private workspaces contain information that is only accessible for the owner of the private workspace. Public workspaces hold knowledge objects that are published via the Internet and accessible by an undefined group of users. Protected workspaces contain knowledge objects that are accessible to a single or a group of knowledge workers that the owner explicitly grants access.

The integration services also support caching of knowledge elements that are accessed repeatedly. A personal knowledge cache contains the knowledge elements of the user's own private knowledge base and of other peers' protected workspaces that the user has access to. The personal knowledge cache is used to optimize network traffic when shortly accessing the same knowledge elements multiple times. The offline cache holds those knowledge elements on the local storage medium that are often accessed by the user while being without a permanent connection to the Internet.

Just as in the centralized case, knowledge and access services build upon the knowledge base. The main difference is that the knowledge repository now is spread across a number of collaborating peers that have granted access to parts of their knowledge repositories. There is no central authority that takes care for the integration of the repositories that participate in a peer-to-peer KMS network. Access and knowledge services are similar to the centralized KMS architecture. However, the peer lacks personalization services as there are no "impersonalized" services in a peer's KMS.

Super peer. In addition to the services offered by a peer, a super peer might access a number of additional, shared data and knowledge sources, e.g., document management systems, content management systems, data warehouses, e-learning platforms, experience data bases or the organization's transaction processing systems. Every super peer consequently extracts, transforms and loads those parts of the data and knowledge sources that fall into the domain handled by a subject matter specialist. Inspection services support the access of documents without the applications that were used to create the documents. The peer-to-peer infrastructure might also provide services for lookup and message handling that improve the efficiency of the distributed KMS.

The integration services offer a shared taxonomy or ontology for the domain handled by the subject matter specialist. This addresses the challenge in a totally distributed KMS that the various knowledge bases cannot be integrated and thus pose a problem for e.g., the interpretation of search results by the knowledge worker. As laid out in Figure B-62, all or a number of subject matter specialists might standardize the terms and meta-data in use and thus provide a common scheme for meta-data, a common taxonomy or ontology for an even larger domain. Super peers might offer replication services to peers that sometimes work offline. Personalization services include portals, profiles and push services that ease the access to the organized collection of (quality approved or even improved) knowledge that a subject matter specialist administers. Access services and knowledge services are the same as the corresponding services of each individual peer.

Peer-to-peer KMS supposedly have the following advantages (Benger 2003, 167f):

- *autonomy*: semi-autonomous organizational units can easily create and share knowledge with the help of those tools and those ontologies that fit their domain,
- *direct communication*: knowledge is exchanged directly without central units that often act as an unwanted filter to knowledge,
- *flexibility*: peer-to-peer KMS allow for the configuration of temporary, dynamic networks of knowledge workers,
- *acceptance*: local storage together with an efficient management of access privileges reduces the barriers to provide knowledge that some central KMS solutions experience.

The peer-to-peer metaphor promises to resolve some of the shortcomings of centralized KMS. Examples are:

- to reduce the substantial costs of the design, implementation and maintenance of centralized KM suites, in terms of hardware, standard software as well as the often underestimated costs of designing, structuring and organizing a centralized knowledge server and the management of users and privileges. This is due to the fact that simple local KMS are often already in place. Compared to a central KMS, additional investments are minimal,
- to reduce the barriers that prevent individual knowledge workers from actively participating and sharing in the benefits of a KMS, e.g., by reducing the psychological barrier to publish knowledge elements to an unknown target group by giving the user full control over the access privileges to her knowledge elements,
- to overcome the limitations of a KMS that (almost) exclusively focuses on organization-internal knowledge whereas many knowledge processes cross organizational boundaries, because workspaces can easily and flexibly be extended to knowledge workers from partner organizations,
- to include individual messaging objects, e.g., emails, instant messaging objects, into the knowledge workspace that are rarely supported by centralized KMS and, moreover,
- to seamlessly integrate the shared knowledge workspace with an individual knowledge worker's personal knowledge workspace.

However, on the other hand, there are still serious technical challenges that have to be overcome in peer-to-peer computing in general. These challenges concern (Barkai 2001, 264ff):

- *connectivity*, e.g., locating peers that do not have public IP addresses and mechanisms for communicating through firewalls,

- *security and privacy*, especially the risk of spreading viruses, unauthorized access to confidential and private information and the installation of unwanted applications,
- *fault-tolerance and availability*, e.g., finding the required resources available when they are needed,
- *scalability*, especially concerning the naming scheme and searches in the flat structure of the distributed search domain,
- *self-managed systems* that are administered by individual users with limited experience and tools who provide services to others and
- *interoperability*, i.e., current peer-to-peer installations cannot connect to each other due to e.g., a variety of computing models, a variety of network settings and a wide range of application types.

There are also a number of organizational issues that still have to be resolved before a peer-to-peer KM infrastructure can be fully deployed in an organization. Examples are (Susarla et al. 2003, 133ff):

- *participation issue*: there have to be incentives to actively participate in the peer-to-peer network in order to foster information sharing and avoid the free rider issue,
- *trust issue*: security and reliability of the peer-to-peer infrastructure have to be believable for the participants of the peer-to-peer network if the system should be used as the sole, personal knowledge workspace of knowledge workers,
- *coordination issue*: structuring and quality management of the knowledge contained in a peer-to-peer network have to be supported in order to avoid information overload.

Working with a peer-to-peer KMS might quickly be less effective and especially less efficient than working with a centralized KMS if the coordinating mechanisms established in a central KMS are missing. Whether actual peer-to-peer solutions will soon overcome the major challenges of a (sufficient!) semantic integration of a variety of heterogeneous knowledge bases, still remains to be seen. Thus, the hybrid architecture proposed here that includes super peers that coordinate parts of the contents and handling of accesses in the KMS might work best for many organizations.

If peer-to-peer KMS are to be successful, they have to address not only the technical and organizational issues, but also have to show how they could resolve the shortcomings of centralized KMS, particularly how a peer-to-peer KMS application system can be seamlessly integrated with the knowledge worker's personal knowledge workspace, what these workspaces should look like, what mechanisms can support the semantic integration of the distributed knowledge workspaces, e.g., a predefined set of dimensions for meta-data, and how working with the peer-to-peer KMS can be made easy enough so that the barriers to participate are not too high. In the following section, Infotop is discussed in detail as an example for a peer-to-peer KMS that also provides ideas on how to address these questions.

7.5.3 Example: Infotop

Infotop[556] is a personal workspace designed to help knowledge workers (1) to organize their personal information and knowledge resources and (2) to share context and collaborate on the basis of peer-to-peer information workspaces. Infotop primarily addresses the challenge of an integrated knowledge workspace for networked knowledge workers. As centralized KMS often only marginally fulfill the requirement of their seamless integration into personal knowledge workspaces, a distributed knowledge environment is found suitable for Infotop. Infotop primarily targets the challenges of accessing, integrating and sharing of knowledge workspaces and proposes to replace the desktop as the primary metaphor for the interaction with personal computers.

Knowledge workers are the primary user group of personal computers. From an ICT infrastructure perspective, the desktop metaphor has been used for decades to administer small amounts of documents. This metaphor has been sufficient as long as the types, formats and amounts of contents to be administered were limited. Today, the desktop provides only a restricted view to the organizational knowledge base. Due to the increase in size and complexity of contents, much of the original desktop's functionality has moved into complex applications, e.g., Web browser, messaging system, document management system, KMS. Thus, the desktop has been replaced in many situations as the central view to collections of contents. This has resulted in today's scenario where there are many applications with many isolated and incompatible views on parts of the data and with many categorizations of these data.

Infotop proposes to replace the desktop with a new metaphor to interact with personal knowledge environments, what formerly was a personal computer. The term Infotop covers the dynamic aspect of knowledge, the flow of knowledge, which is best described by the term information. Infotop thus means to be "on top of the information" that flows in and out of the personal knowledge environment.

Desktop metaphor. A metaphor is one thing conceived as representing another. Using metaphors takes advantage of peoples' knowledge about them. For example, people in offices have been used to store paper documents in file folders. It makes sense to these people to store computer documents in folders on the computer, i.e., in containers that look and behave like folders. The desktop is the primary metaphor being used as interface on personal computers. It was introduced when computers were quite different to today's machines, see (Genter/Nielson 1996). While computers, users and the environment have changed, interfaces and the basic handling of data have stayed the same[557]. The desktop has become an unmanageable mess (Tristram 2001). Countless files are stored on increasingly more capacious

556. This section summarizes joint work done by the author and Sametinger which has been presented e.g., in Maier/Sametinger 2002, 2003, 2004.
557. See also section 4.1.3 - "From traditional work to knowledge work" on page 46.

storage drives. This has resulted in big hierarchies of folders that make it difficult to retrieve information. The problems can be summarized as follows:

- storing contents on traditional desktop and folder systems is limited to one single hierarchical folder structure rather than a flexible means of categorization,
- there are trivial and multiple categorization mechanisms in various applications, e.g., folder structure, personal information management, email system, Web browser,
- meta-data and versioning data are only available with specific applications,
- multiple documents are different representations of the same contents, e.g., a text document in the format of the text processing system, postscript and the portable document format, and
- there is insufficient meta-data about local and remote documents.

Infotop. Rather than having a desktop with a hierarchic folder view, Infotop supports multiple views on documents and a much more powerful way of accessing information. Two perspectives have to be considered.

- *island approach*: can be applied to a single computer and a single knowledge worker (a single peer). This computer may be connected to other machines, but there is no extra communication in support of the island approach.
- *peer-to-peer approach*: comprises many knowledge workers who use Infotop and, thus, can benefit from advanced features and shared context when communicating and working together.

Downward compatibility is a necessity in order to consider a shift to the proposed approach. Therefore, today's desktop metaphor with files and folders should be a special case or view of Infotop. Subsequently, the island approach is described with Infotop's six *dimensions* for the categorization and visualization of knowledge. Due to its importance, the dimension *time* requires extra consideration. Also, *multi-dimensional views* and the handling of *meta-data* are described. Then, the peer-to-peer approach is shown, especially the *shared context* of collaborating users, the support of *knowledge work processes*, the proposed *peer-to-peer architecture*, and some thoughts about Infotop's *implementation*.

Dimensions. Business intelligence software allows users to quickly analyze data that has been transformed into a subject-oriented, multidimensional data warehouse (Inmon 1992). Online analytical processing (OLAP) tools are used to perform trend analysis and statistics on e.g., sales and financial information in an interactive question-answer way. Infotop uses the six dimensions time, topic, location, person, process and type for effective categorization, visualization and navigation of collections of contents. In analogy to OLAP techniques, these dimensions are used for slicing, dicing, drilling down, rolling up, and ranging operations on contents of a personal knowledge environment:

- *time*: any representations with a timed order,
- *topic*: any topics a user is interested in,

- *location*: any geographic location like a city or country; local vs. LAN vs. Web,
- *person*: any person, physical or not, e.g., a company, an organizational unit,
- *process*: any project or process, e.g., a conference, a paper writing process, an administrative task with many steps,
- *type*: any type of document, e.g., text document, MS Word document audio or video file.

Figure B-63 shows a simple one-dimensional view where documents are shown that belong to various topics. On the right-hand side, there are six buttons that can be used to switch to different dimensions and to select sets of documents that are displayed in these dimensions. The pile metaphor (Mander et al. 1992) can be used to display information about sets of documents. Additionally, the numbers of documents are indicated for each displayed topic.

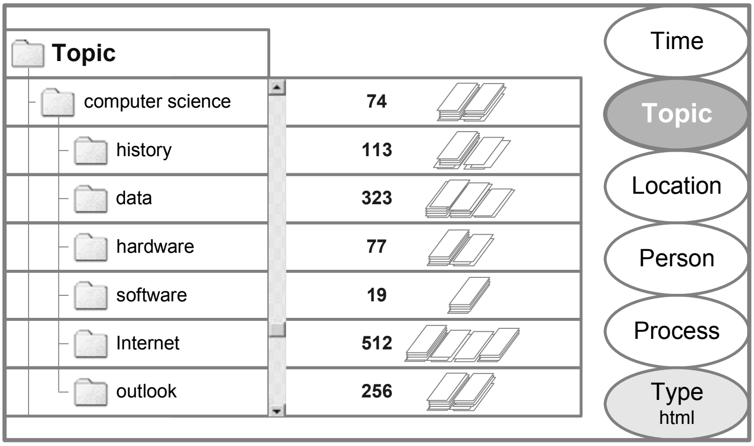

FIGURE B-63. Infotop - one-dimensional view

Visualization techniques like the well-known icons, thumbnails or lists are useful when displaying sets of documents. The knowledge worker can arbitrarily define several hierarchies of any of these dimensions and use them for display, e.g., in case of the dimension person the three hierarchies author, sender, receiver. One simple hierarchy for topics can be seen in Figure B-63. Views may be restricted to documents with specific attributes, e.g., documents of a specific process or documents of a specific range of dates. In Figure B-63, only documents of type html are displayed.

In addition to the dimensions, OLAP tools present facts in selected cells of a resulting spreadsheet, e.g., the amount of products ordered according to the dimensions *customer* and *region*. In this case, facts represent information on sets of contents represented in each cell, could be for example:

- the number of elements as represented in Figure B-63,
- the amount of data, e.g., the number of pages or MBytes used,

- the number of contributions or of questions answered of knowledge providers,
- an aggregate valuation of elements, e.g., the number of accesses to elements,
- a measure of the skill levels of knowledge providers in a domain, or, in finer granularity,
- any other meta-information that is stored along with elements, e.g., the titles of documents, or
- a comparative measure, e.g., the proximity of competencies between a number of potential knowledge providers in a certain domain.

Time. Time is one of the most crucial attributes of documents, e.g., time of creation, time of last modification, time of last read only access. Usage statistics may also be useful, such that frequently used documents can stand out. Figure B-64 shows documents assigned to the topic knowledge management that have a relation with the ECKM 2002[558] conference in a calendar view. The time of last modification is considered for display. Clicking one of the days will bring up information about all documents, i.e., icons or a list with detailed information.

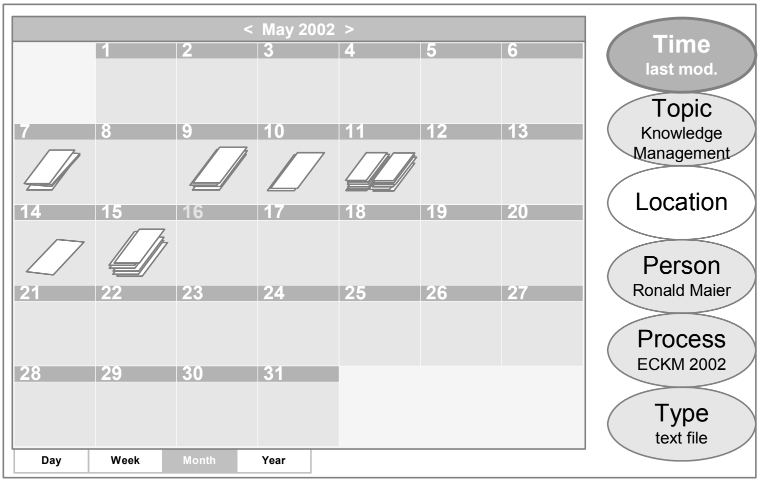

FIGURE B-64. Infotop - time-oriented view

Apart from usual appointments it is useful to have email messages, text documents and other forms of documents, e.g., comments, yellow stickers, displayed in calendars. It is also useful to display a selection of documents, e.g., all documents related to a project displayed in the calendar, or all documents of a person, i.e., all email messages from and to that person, all files exchanged with that person, all Web documents about that person that the knowledge worker has visited, etc.

558. European Conference on Knowledge Management 2002.

Multi-dimensional views. OLAP tools help users to interact with multitudes of statistics in order to isolate specific items. Infotop supports similar mechanisms to browse, navigate and filter information. The hierarchies can be used for this purpose. For example, the knowledge worker can select the two dimensions *process* and *person* for viewing, see Figure B-65.

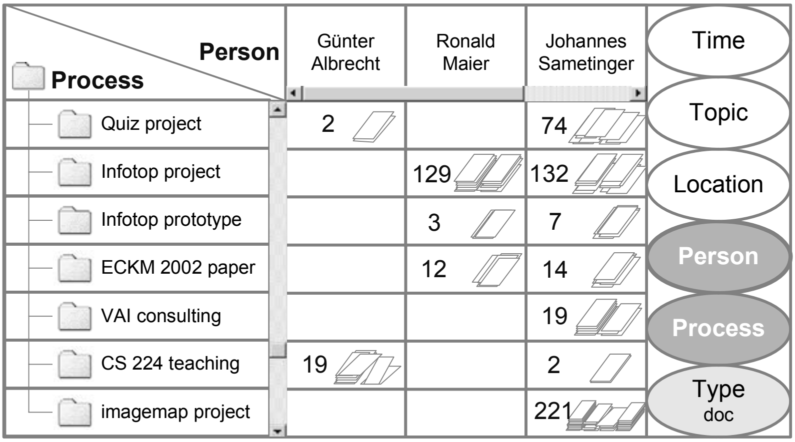

FIGURE B-65. Infotop - two-dimensional view

Six dimensions enable to select documents in one hierarchy and display this selection in another hierarchy. For example, select all Austrian documents, i.e., documents with location= Vienna, location= Linz, or location= any other Austrian location, and then display the documents according to a hierarchy based on persons.

Meta-Data. For efficient document retrieval and for grouping of documents, categories have to be associated with documents. Attributes have to be assigned with documents. This can become a nuisance to the knowledge workers, because they may not want to manually categorize each incoming and outgoing email message, or each Web page that they have visited. Therefore, an automated, or at least a semi-automated approach is needed for this task. A couple of attributes should be defined for each document, e.g., title, author, date, event, location, person, process. Each attribute of a document has an undefined or a defined value, e.g., location= Dublin, date= 9/25/2002. The meta-data can easily be extracted from context that comes with a document or the activities that are performed on a document, e.g., in the case of an email message Infotop can derive sender, receiver (person, location), date (time), subject (topic, process) and type of attached file (type).

Shared context. Users have information on their private computers and can also access public resources, typically on the Internet. Additionally, servers on local area networks provide extra information that is not accessible to the public, but only to a restricted number of users. Infotop separates a private, a protected and a public workspace. Private workspaces contain information that is stored locally on each knowledge worker's computer and accessible only for the owner of the private workspace. Public workspaces include information that is published via the Internet and accessible by an undefined group of users. Protected workspaces lie somewhere in between. They contain information that is not accessible for everyone, but for whoever the owner grants explicit access, e.g., digital libraries.

Private, protected and public workspaces of an individual knowledge worker can be placed on that worker's personal computer, see user 3 in Figure B-66. Additionally, user 3 shares in parts of other users' workspaces. The dashed line and the gray boxes indicate the shared-context information workspace of user 3, i.e., a virtual workspace that includes user 3's private, protected and public workspace as well as all public and parts of protected workspaces of other users. It is important to note that a user's protected workspace is not open to the public, but rather allows restricted access only to those individuals that the user wishes. Thus, access privileges of the protected workspace have to be configurable in a flexible manner. Typically, public workspaces grant permission to read only, whereas protected workspaces may be open to write.

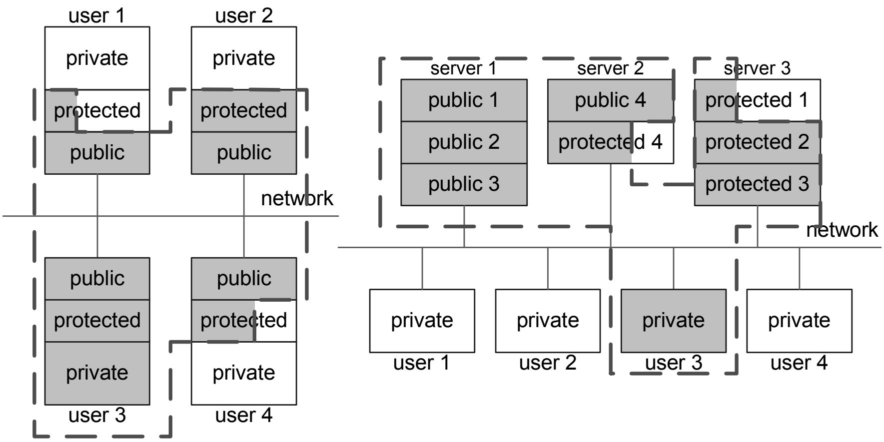

FIGURE B-66. Infotop - alternative architectures

The peer-to-peer approach on the left-hand side of Figure B-66 is contrasted with a client-/server-approach on the right-hand side. Infotop's concepts work in both worlds. In the server architecture, only private workspaces reside on the individual users' personal computers whereas the public and protected workspaces are submitted to dedicated servers. This architecture resembles most to the centralized KMS architecture as described in section 7.3.3, p. 311ff. In order to achieve the benefits promised by the peer-to-peer metaphor, the Infotop approach institutionalizes private, protected and public workspaces on all workplaces (Figure B-66, left-hand side). Additionally, any information in these workspaces has to have meta-information attached, according to the six dimensions mentioned above, such that powerful query mechanisms can be supported. Assignment to e.g., topics is crucial for workspaces. This supports several virtual workspaces for different topics of interest, i.e., several dashed lines in Figure B-66. Virtual workspaces can overlap, because workspaces and sets of documents can be assigned to more than one topic.

Organizing and visualizing this shared-context information workspace for each individual remains a challenging task. In the following, the multi-dimensional workspace as described above can be used with minor modifications in a shared context (see Figure B-67).

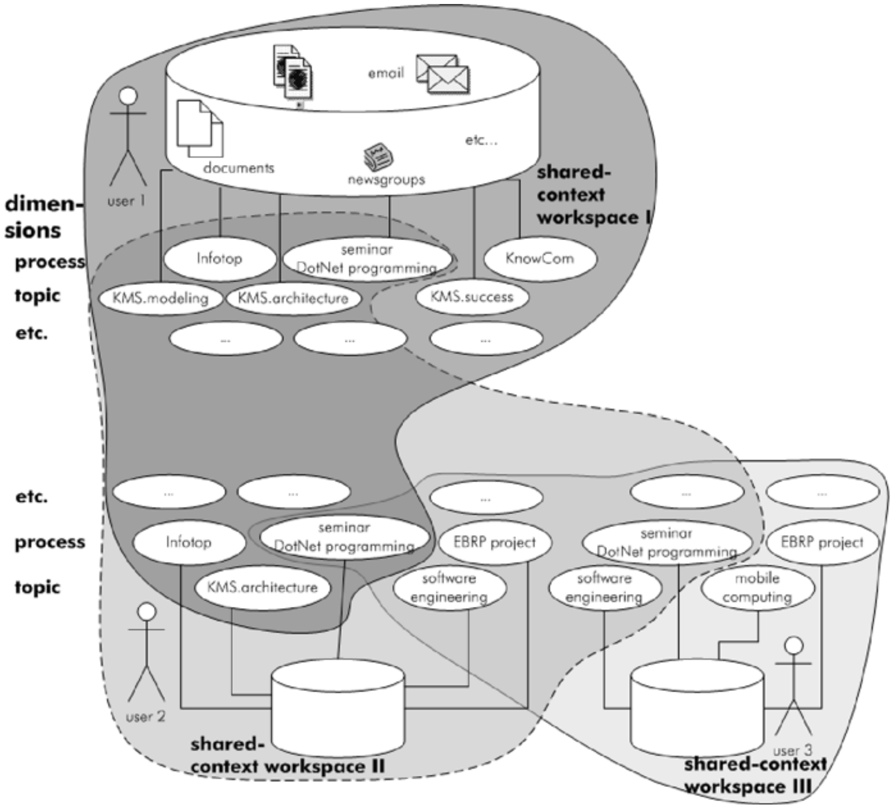

FIGURE B-67. Example for shared-context workspaces in Infotop

Figure B-67 shows how the dimensions of Infotop can be used to define shared-context workspaces and, thus, to distinguish private from protected information. Users 1, 2 and 3 all have access to their personal data store that is visualized by the data base symbol. The data store can contain text documents, personal information management documents, e.g., addresses, bookmarks, calendar with appointments, to-do-lists, hypertext documents, messaging objects, such as emails, contributions to newsgroups, multimedia elements, etc. Infotop provides access to the entire personal data store using its six dimensions. In Figure B-67, the two dimensions process and topic are used to define shared-context knowledge workspaces. User 2 grants user 1 access to all data in her data store that are assigned to “Infotop” and “Seminar DotNet Programming” in the process dimension and all data assigned to “KMS architecture” in the topic dimension whereas the “EBRP project” and the topic “software engineering” are not accessible to user 1. User 1 grants user 2 access to all data in his data store that are assigned to “Infotop” and “Seminar DotNet Programming” in the process dimension and all data assigned to “KMS modeling” and “KMS architecture” in the topic dimension whereas the “KnowCom” process and the topic “KMS success” are not accessible for user 2. Consequently, workspace management is easily accomplished in a flexible manner by assigning instances of each of the six dimensions to (groups of) users.

The six dimensions are helpful, no matter whether the information is private or shared. They have been introduced to get rid of the rigid file hierarchy. The shared context should conceal network structures and stress the logical boundaries among knowledge elements. However, explicit consideration of workspaces and thus a seventh dimension may be necessary to visualize social networks and promote the sharing of context.

Knowledge work processes. Figure B-68 outlines how Infotop supports important knowledge work processes. A user externalizes, distributes, submits, acquires, searches, applies information in her shared-context information workspace. The solid ellipse in Figure B-68 depicts the user’s individual workspace, while the dotted ellipse depicts the user’s shared-context information workspace.

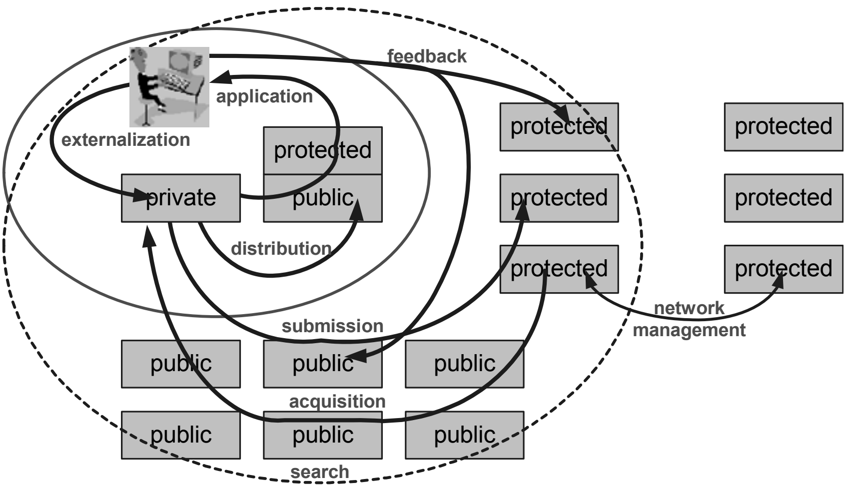

FIGURE B-68. Infotop - knowledge work processes

- *Externalization process*. Externalization of information is done with regular applications, e.g., a word-processor, or (co-)authoring tools. This process results in documents that typically are at first stored in the private workspace. It is important to have meta-information attached to these documents. This is not sufficiently supported by today's applications. Infotop provides rich contextualization of documents using the six dimensions.

- *Submission process*. In the simplest case, submission means publication of a new knowledge element and its distribution towards a topic-oriented network, i.e., in a protected or public workspace. Versioning of information and the support of workflows is required for the submission process.
- *Distribution process*. The distribution process involves moving or copying information from one’s private to one’s protected or public workspace. It is useful to have this process combined with some sort of notification, especially in the protected workspace.
- *Search process*. Searching is done primarily based on meta-information in one’s workspace consisting of one’s private, accessible protected and public workspaces. Protected and public workspaces have to be prioritized according to topics, e.g., workspaces of research groups have to be considered only when the search process is aimed towards the research topics of these groups. Findings in protected workspaces are typically more relevant than findings in public workspaces.
- *Application process*. The application process involves any usage of information that has been retrieved from an arbitrary source, i.e., from protected and/or public workspaces.
- *Feedback and improvement process*. Responses or reflections to information in an arbitrary workspace can improve the quality of information. Feedback includes communication to information holders, i.e., workspace owners, citations, etc.
- *Acquisition process*. The acquisition of information includes the extension of the search domain to include new workspaces, the location of information in any of the accessible workspaces and copying this information or a link to it into one’s individual workspace.
- *Community or network management process*. Communities[559] share their interest in certain topics. It is necessary to have topic directories in public work spaces, where users can register and obtain permission to participate in protected workspaces that are assigned to these topics. The consideration of new topics results in new dashed lines, see Figure B-66. The acquisition of information is supported by the extension of one’s workspace by including additional protected workspaces.

Figure B-69 shows how the knowledge work processes discussed above can be supported in a setting with a number of knowledge workers using Infotop and collaborating in overlapping knowledge communities (see also Maier/Hädrich 2006). In Figure B-69, three communities are visualized. Communities correspond to

559. See section 6.1.3 - “Groups, teams and communities” on page 177.

shared-context workspaces in which a number of knowledge workers participate. In the terms of Figure B-67, externalization of knowledge requires documentation of a knowledge element, organization according to the six dimensions and moving it into Infotop's knowledge base. Submission simply means that access privileges are granted to members of a community for instances of one or more Infotop dimensions. The search domain used in a search process consists of all locatable peers that have granted access to their knowledge base. Priority is given to those peers that participate in the same community the topic of which most closely matches the search term.

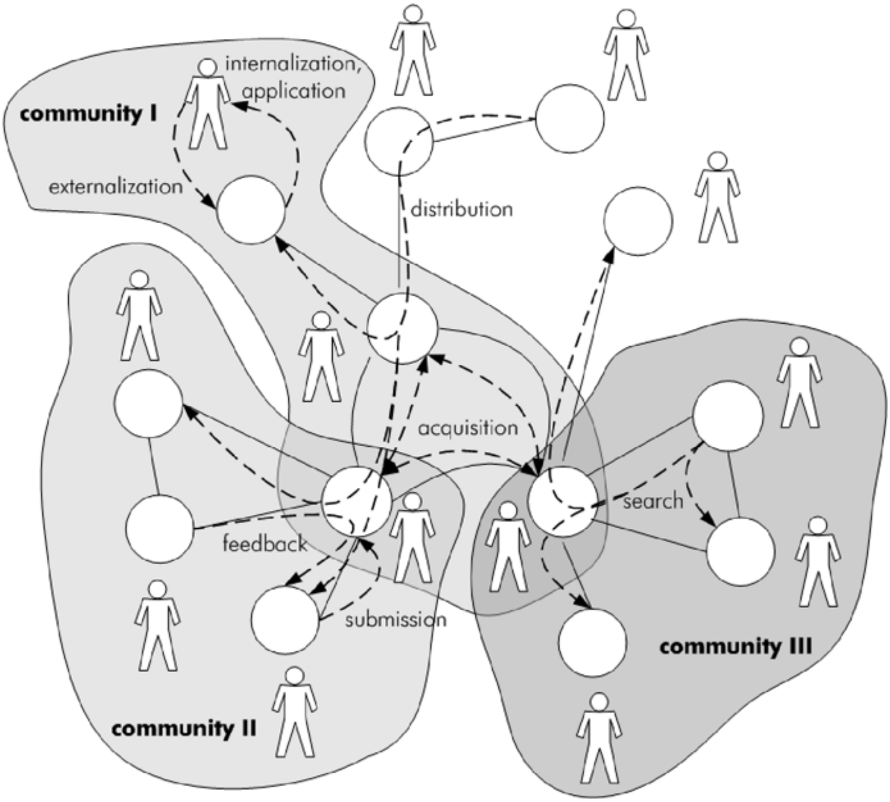

FIGURE B-69. Knowledge work processes supported by Infotop[560]

Peer-to-peer architecture. Infotop addresses all three main issues, (1) participation, (2) coordination and (3) trust, that challenge peer-to-peer KMS as identified above[561]. Participation should be no more of a problem than in centralized KMS

560. See also Maier/Sametinger 2004 for a preliminary version of this figure.
561. See section 7.5.2 - "Peer-to-peer knowledge management systems" on page 342.

within organizational boundaries. Moreover, if Infotop can provide a useful solution to personal knowledge management that does not require any additional effort to establish shared workspaces in a peer-to-peer network, a large number of users might be convinced to participate. In peer-to-peer knowledge networks that cross organizational boundaries, (professional) communities along with personal contacts, contracts, shared goals and interests might act as a kind of social infrastructure that induces social regulations and also trust into the peer-to-peer network.

Figure B-70 shows the architecture of one Infotop peer that consists of the four layers infrastructure, integration, knowledge and access services.

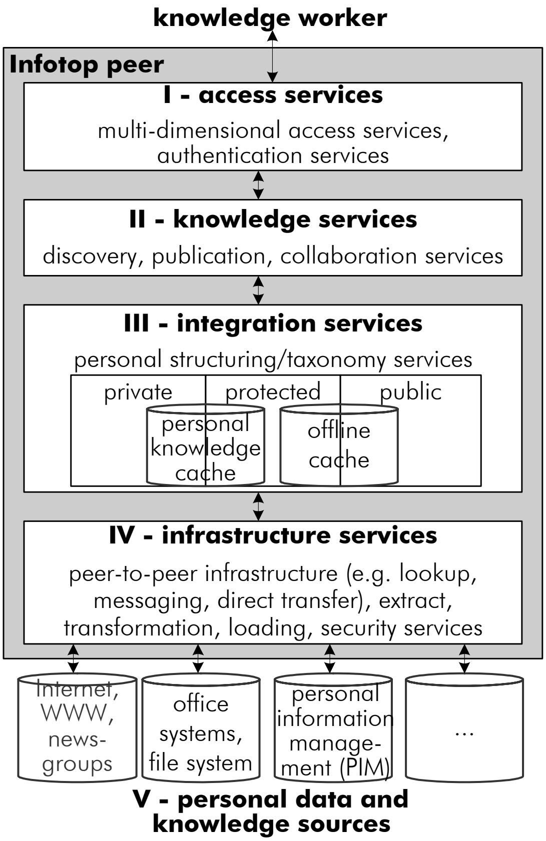

FIGURE B-70. Architecture of one Infotop peer

The architecture is closely tied to the ideal peer-to-peer KMS architecture and therefore includes the same layers as the centralized architecture[562], but lacks a

562. See section 7.3.3 - "Integrating architectures for KMS" on page 311 and section 7.5.2 - "Peer-to-peer knowledge management systems" on page 342.

centralized knowledge structure, taxonomy and repository. Personal data and knowledge sources are extracted, transformed and loaded into an integrated Infotop knowledge base. The integrated knowledge base comprises a private, protected and public area. A personal knowledge cache is used to optimize network traffic when shortly accessing the same knowledge elements multiple times. Due to the fact that knowledge workers might still at some time prefer to work offline, this knowledge base has an offline cache keeping those knowledge elements that are often needed on the local storage medium preferred by the knowledge worker. Just as in the centralized case, knowledge and access services build upon this integrated knowledge repository. The main difference is that the knowledge repository now is spread across a number of collaborating peers that have granted access to parts of their knowledge repositories. As Infotop itself realizes an integrated knowledge workspace, there are no translation and transformation services in the access layer. Instead, the access layer provides Infotop's visualization concept with the six dimensions time, topic, location, process, person and type as well as the OLAP-type functions as the main interface to collections of contents, both personal and shared across multiple workspaces of networked knowledge workers.

Implementation. Currently, Infotop's concepts are improved and implemented as a joint effort by two work groups headed by Sametinger at the Johannes-Kepler-University Linz[563] and by the author at the Universiy of Innsbruck. The implementation is based on a combination of Web services, data base, peer-to-peer and configuration management technology. Web services and peer-to-peer-technology can be used to seamlessly integrate other users' shared workspaces into one's own workspace in a platform-independent way. A data base is required in order to manage the meta-data created by Infotop. Configuration management and version control is needed to avoid versioning conflicts and to allow coordinated and cooperative work in the shared context. Also, Infotop has to exchange meta-data with other applications, e.g., messaging, office management and a search engine. The presentation of the workspace has to be modeled according to Infotop's six dimensions.

To sum up, KMS are typically restricted to one organization's boundaries. A significant portion of knowledge work processes crosses these boundaries and thus can only be supported on the level of a personal knowledge workspace. Infotop should act as the main access point both for personal knowledge management and for ad-hoc collaboration in a shared context. It is important to include multiple ways to visualize the structure of elements in the dimensions, such as hierarchies, networks (knowledge maps) and geographical information systems in order to meet individual visualization needs. Another promising direction for Infotop is to integrate personal KM techniques, e.g., portfolios, visualization of individual knowledge workers' knowledge status, learning and networking needs, with corporate KM instruments, e.g., content management, yellow pages, communities,

563. URL: http://www.se.jku.at/sametinger/.

project staffing or competence development programs. Infotop plays the role of an enabler and catalyst to spark usage of corporate KMS solutions and start a positive, reinforcing cycle of more and more active, motivated participants handling knowledge in organizations.

7.6 Classification

There are a great number of information and communication technologies that are discussed as supporting knowledge management. Apart from more traditional tools and systems as discussed in this book as the technological roots of KMS (section 7.1) there are a great number of functions providing knowledge-related services. These services have been combined into a centralized KMS architecture (sections 7.3.3, 7.4). As a contrast, section 7.5 has shown an alternative way of organizing KMS, a decentralized, peer-to-peer architecture.

However, both architectures can be seen as ideal in the sense that almost all actual tools and systems offered on the market or implemented in organizations only offer a certain portion of these services. The following section aims at organizing the abundant number of tools and systems that are discussed as being helpful for KM. Firstly, a number of classifications of tools and systems in support of KM as found in the literature are presented (section 7.6.1). The tools are then ordered into a classification scheme (section 7.6.2).

7.6.1 Knowledge Tools

There are a great number of tools, platforms and application systems on the market which claim support for organizational memory or knowledge management respectively[564]. The field is still immature, though, in the sense that there are no classes of systems that the literature has agreed on. So far, there are several proposals for classifications of systems which mostly lack completeness and also exclusiveness in the sense that one system fits into one and only one category. Table B-19 shows a comprehensive overview of classifications of technologies, tools and systems supporting KM[565]. Classifications in the literature fall into two categories:

Market view. These classifications try to cover either *technologies, tools and systems that potentially support KM* (wide view) or they cover the *functionality of KMS* (narrow view).

Theoretical view. These classifications are based on existing models describing *types of knowledge* (abstract view) or *KM, OL or OM processes or tasks* respectively (concrete view) that could potentially be supported by ICT in general or KMS in particular.

---

564. For a list of KMS see the support Web site for this book http://iwi.uibk.ac.at/maier/kms/ .

565. See also Maier/Klosa 1999c, 8ff, Klosa 2001, 63ff for a detailed discussion of some of the classifications listed here.

TABLE B-19. Classifications of technologies, tools and systems supporting knowledge management

author(s)	categories
classifications on the basis of types of knowledge	
Schüppel 1996, ILOI 1997 propose a classification according to the knowledge supported by KMS	• explicit—implicit knowledge • current—future knowledge • internal—external knowledge • experience-based—rationality knowledge
Warschat et al. 1999, 55ff classify ICT to support KM using the hierarchy symbol, data, information, knowledge	• data warehouse systems • document management systems • Web publishing systems • content management systems • knowledge-based information systems
classifications on the basis of knowledge management tasks, life cycle or strategies	
Apostolou/Mentzas (1998, 3.3) use Nonaka's (1991, 98f, 1994, 18f) knowledge transfer processes	• socialization (e.g., email, discussion lists, bulletin board, multimedia conferencing) • internalization (e.g., lessons learned DB, hypermedia CBT, process-history tracking, data warehouses, data mining) • externalization (e.g., semantic networks, ontologies, push technologies, agent technologies, issue-based argumentation, data warehousing) • combination (e.g., document management systems, workflow management systems, group decision support systems, search and filtering systems, computer-mediated communication)
Dieng et al. (1998, 3ff) classify methods and tools according to their support for phases of corporate memory management	• detection of needs • construction of the corporate memory • diffusion of adequate elements of the corporate memory • use of the corporate memory • evaluation of the corporate memory • maintenance and evolution of the corporate memory
Mentzas et al. (2001, 95f) classify KM software using two dimensions reflecting Hansen et al.'s (1999) two KM strategies: process-centered versus product centred.	• primarily process-centered (knowledge transfer, personalization): shared files, email, real-time messaging, net conferencing, discussion groups, white-boarding • primarily product-centered (knowledge content, codification): file management systems, full text retrieval, structured document repositories, semantic analysis, knowledge maps, Intranet • about equally high on both dimensions: push technology, automatic profiling

TABLE B-19. Classifications of technologies, tools and systems supporting knowledge management

author(s)	**categories**
classifications on the basis of definitions and models of OL and OM	
Ackerman 1994 classifies ICT to support the organizational memory	• store and retrieve data (e.g., data base systems) • share and publish information • OM and group memory components (new forms of systems, e.g., Answer garden, group meeting systems) • capture design processes and informal communication • access members of the organization • develop knowledge structures • document management systems • platforms (e.g., Lotus Notes)
Jacobsen 1996, 169 classifies ICT according to the two dimensions acquisition and deployment of competence	• acquisition of competence: perception, learning, application • deployment of competence: transfer, storing
Stein/Zwass (1995, 97ff) propose a framework for organizational memory information systems which consists of two layers (see also section 4.3 - "Knowledge management systems" on page 82)	layer 1: competing values model of effectiveness: functions of organizational effectiveness • integrative subsystem • adaptive subsystem • goal attainment subsystem • pattern maintenance subsystem layer 2: information processing model of memory: mnemonic functions • knowledge acquisition • knowledge retention • knowledge maintenance • knowledge search and retrieval
classifications on the basis of the functionality of KMS	
The **Delphi Group** (1997, 14) suggests five groups of KMS functions reflecting a narrow focus on explicit, documented knowledge: a knowledge repository, and a set of tools to filter, organize and present this knowledge (Delphi 1997, 15).	• intermediation: brokering information or knowledge seekers and knowledge providers • externalization: capturing knowledge, structuring and organizing it in a repository according to a framework or ontology • internalization: extraction and filtering of knowledge from a repository • cognition: system functions to make decisions based on available knowledge • measurement: measure, map and quantify corporate knowledge and the performance of KM solutions

TABLE B-19. Classifications of technologies, tools and systems supporting knowledge management

author(s)	categories
Apostolou/Mentzas (1998, 3.4) define ICT services which are part of a knowledge leveraging infrastructure	• knowledge search, retrieval and navigation • knowledge indexing, mapping and classification • knowledge storage, analysis and meta-data processing • knowledge distribution and publication • collaboration
Borghoff/Pareschi 1998a, 5f classify ICT specifically supporting KM. The classification is rooted in an empirical study of IT requirements for KM done by **Xerox** which in turn is based on Nonaka's (1991, 98f, 1994, 18f) knowledge transfer processes[a].	• knowledge repositories and libraries (documents): search, heterogeneous document repository, access, integration and management, directory and links, publishing and documentation support • communities of knowledge workers (people): awareness services, context capture and access, shared workspace, knowledge work process support, experience capture • flow of knowledge: using knowledge, competencies and interest maps to distribute documents to people • knowledge cartography (navigation, mapping and simulation): tools to map communities of practice, work process simulation, domain-specific concept maps, maps of people's competencies and interests (yellow pages), design and decision rationale
Bair (1998, 2) identifies four dimensions of functionality that differentiate KM technology from other (software) products	• semantic functionality: extends document and content management to increase the relevance of retrieved/pushed information and handle dynamic semantic relationships: categorization of documents, semantic networks, natural language processing • collaborative functionality: builds on Groupware, email and workflow technology to support the capturing of (tacit) knowledge: identification of experts based on skills, recognition, publications; collaborative filtering (e.g., evaluation of documents) • visualization functionality: use advanced graphical techniques to display relationships between knowledge elements • scale/integration: the ideal system provides access to all information resources in the organization as well as external resources, to any data type and to any application, including data warehouses
Wiemann (1998, 7ff) classifies systems for KM according to their impact on the knowledge	• divergent systems: support knowledge exchange between employees with no attention to quality assurance or synthesis of knowledge elements/contributions, e.g., communication systems, platforms for document exchange, skills data bases • convergent systems: systematically identify, evaluate, document, refine, categorize and provide access to knowledge elements, e.g., in the form of a data base of best practices

TABLE B-19. Classifications of technologies, tools and systems supporting knowledge management

author(s)	**categories**
Zack (1999a, 50), distinguishes systems supporting the handling of explicated knowledge and systems supporting collaboration of experts	• integrative KMS: electronic publishing, integrated knowledge base • interactive KMS: distributed learning, forums
Meso/Smith (2000, 227ff) identify ten key technologies for organizational KMS which support the functions using, finding, creating and packaging knowledge	• computer-mediated collaboration • electronic task management • messaging • video conferencing and visualization • group decision support • Web browsing • data mining • search & retrieval • intelligent agents • document management
Seifried/Eppler (2000, 31ff) define a KM suite as an open IT-platform that integrates four function areas	• collaboration: computer-supported cooperative work, computer-supported cooperative learning, workflow management • content management: document management, personal information management, group information management • visualization & aggregation: knowledge maps, knowledge portals, taxonomies, directory services • information retrieval: search methods, search results, search languages, sorting, retrieval
Versteegen (2000, 101) categorizes tools for KM according to their focus on bundles of KMS functions	• modeling and analysis of knowledge • storing and administration of knowledge • distribution of knowledge • access to and retrieval of knowledge
Alavi/Leidner (2001, 114) distinguish common applications of IT to KM initiatives	• coding and sharing of best practices • creation of corporate knowledge directories • creation of knowledge networks
Jackson (2001, 5f) classifies systems for gathering, dissemination, synthesis, communication and storage of knowledge	• document management systems • information management tools • searching and indexing • communications and collaboration • expert systems • systems for managing intellectual assets: mostly legal systems to maintain trademarks, patents and other intellectual property

TABLE B-19. Classifications of technologies, tools and systems supporting knowledge management

author(s)	categories
classifications of technologies supporting knowledge management	
Schmoldt/Rauscher (1994) classify technologies for KM which are rooted in AI	• knowledge-based systems • visualization systems • virtual reality systems • spatial data management: management of data using spatial models and spatial display techniques • computer-supported cooperative work / Groupware • hypertext systems
Allee (1997, 224f) identifies basic technologies as "musts" for knowledge-based organizations	• document management • on-line access: to documents, data bases for every employee • email connectivity: expanding simple communication to topic-based information resources, conferencing and bulletin boards • expert systems: for decision making and performance support • pattern-recognition: e.g., data, text mining, knowledge discovery in data bases
Allweyer (1998, 40ff) uses Scheer's (1998) four-level architecture of business process management for the classification of technologies supporting the management of knowledge processes[b]	• design level (modeling and analysis of knowledge processing, knowledge process re-design): tools for modeling, documentation, analysis and navigation of knowledge processes • management level (performing, controlling, monitoring, improvement of knowledge processes): tools and functions supporting controlling and monitoring of knowledge processes • steering level (distribution and sharing of knowledge, search of and access to knowledge): Groupware, Intranet, search engines • application level (creation, documentation, application of knowledge): office information systems, CAD, data bases, knowledge-based systems
Ruggles (1998, 82ff) surveys organizations and classifies the technologies that are implemented as part of a KM initiative	• Intranet: create an Intranet with KM in mind, e.g., for sharing information between (virtual) teams • knowledge repository: develop knowledge repositories and data warehouses to capture explicit, codified, contextualized knowledge • decision support tools: focus is on managerial decision making • collaboration: implementing groupware to support groups in generating, structuring and sharing of knowledge

TABLE B-19. Classifications of technologies, tools and systems supporting knowledge management

author(s)	categories
In the **Competence Center Business Knowledge Management** in St. Gallen several classifications of systems supporting KM have been developed (Thiesse/Bach 1999, 91ff, also IWI HSG 1998, 22)	• search engines: with the functions crawling, indexing, ranking, searching (e.g., Verity) • data warehouse/business intelligence systems • workflow management systems • document management systems • Web management systems (e.g., Gauss, Intranetics) • push services (e.g., GrapeVine, ChannelManager) • knowledge mapping (e.g., Aptex) • multimedia bases (e.g., InXight) • Intranet platforms • other tools for integration and information retrieval
Astleitner/Schinagl (2000, 173ff) classify software tools that are relevant for KM	• tools for information retrieval • data bases • broad KM tools (e.g., Lotus Notes, Grapevine) • focused KM tools (expert systems, constraint-based systems) • real-time KM tools (case-based reasoning) • long-term analysis tools (neural nets) • search engines • intelligent agents • Groupware • integrated performance support systems • tools for data mining and data analysis
Binney (2001, 37ff) reviews KM-enabling technologies classified according to six categories of theoretical KM approaches	• transactional KM: (rule-based) expert systems, cognitive technologies, semantic networks, probability networks, rule induction, decision trees, geo-spatial information systems • analytical KM: intelligent agents, Web crawlers, DBMS, neural computing, push technologies, data analysis and reporting tools • asset management KM: document management tools, search engines, knowledge maps, library systems • process-based KM: workflow management, process modeling tools • developmental KM: computer-based training, on-line training • innovation and creation KM: groupware, email, chat rooms, video conferencing, search engines, voice mail, bulletin boards, push technologies, simulation technologies

TABLE B-19. Classifications of technologies, tools and systems supporting knowledge management

author(s)	categories
Hoffmann (2001, 78f) gives a broad classification of basic technologies used in KM solutions	• Intranet technology • Groupware • electronic document management • information retrieval tools • workflow management system • data analysis and data warehousing • agent technology • help desks • machine learning • computer based training
IBM (Tkach 2001) identifies five strategy areas for KM efforts for which there are already existing tools and application systems	• business intelligence: analyzing data bases using data mining, data warehousing and OLAP, focusing on explicit knowledge • collaboration: expert modeling and decision-making analysis, focusing on tacit knowledge • knowledge transfer: identifying and launching communities or virtual teams, distributed and distance learning technology • knowledge discovery and mapping: text mining techniques and techniques for the contextualized representation of knowledge sources (people and information), clustering, classification and visualization of documents • expertise: organizational network analysis, expert yellow pages and networks; finding, cataloging and making available the best expertise in organization

a. See also Böhmann/Krcmar 1999 who use the same Xerox classification, but propose an extended list of ICT supporting KM, e.g., computer-supported cooperative learning, collaborative filtering, enterprise information portals, meeting support systems.

b. The term *knowledge process* as used by Allweyer (1998, 44) denotes knowledge-intensive business processes in the terminology used in this book (see section 6.3.2 - “Knowledge management processes” on page 212)

There are almost as many different classifications for KMS as there are authors who tried to shed some light on this still diffuse market. Taken together, the classifications comprise a wide field of tools and systems which basically fall into one of the following four categories:

Technological roots of KMS. This group comprises more traditional ICT which can be used to support KM initiatives, such as Groupware, data warehouses, business intelligence tools, modeling software, document management systems, workflow management systems[566].

566. For a brief description see section 4.3 - “Knowledge management systems” on page 82, a detailed description can be found in the literature on each of these technology bundles.

ICT platforms. Corporate Intranet infrastructures or Groupware platforms such as Lotus Notes can be designed "with knowledge management in mind"[567].

Specialized KM tools. Some KM tools have roots in the AI field and perform specific functions necessary for KM. Other KM tools are necessary to integrate several of these functions or several of the more traditional ICT[568].

KMS in a narrower sense. These comprehensive, integrated KMS solutions are also called KM suites and integrate a large set of technologies for knowledge sharing under a common platform.

In the following, the focus will be on ICT bundle solutions or platforms and therefore KMS in a narrower sense as a detailed description of each of the other bundles could fill bookshelves of literature.

7.6.2 Classes

In addition to the theory-driven classifications in the literature (see Table B-19), the theory-driven approaches can be based e.g., on the dimensions

- *contents* of the systems with respect to the types of knowledge focused,
- *roles* of the users, e.g., participants, knowledge broker, knowledge manager, members of the organization vs. externals,
- *organizational level*, individual, collective, e.g., group, team, community, entire organization as well as
- the *technologies* used.

The market-driven classifications vary basically with respect to how narrow or wide the focus on KM-related technology is.

In the following, a *market-driven classification* with a narrow focus on specialized tools for KM, or KMS, is presented which is based on bundles of functions of KMS[569]. The classification of KMS results from a detailed market survey of KM tools and systems conducted by the author (Maier/Klosa 1999c). Then a pragmatic theory-driven classification will be presented which will be used in part D to support different KM scenarios.

The KMS or KM suites are operated on the basis of an (organization-wide) information and communication infrastructure, in most cases an Intranet platform or Lotus Notes environment, on which information sharing between (virtual) teams, both within the organization and across organizational boundaries with allies, suppliers and customers is possible. The basic functionality of such an ICT platform designed "with knowledge management in mind" would comprise an integrated set of the following bundles of functions:

---

567. See section 4.3 - "Knowledge management systems" on page 82.
568. For a brief description of some examples of this category see section 4.3 - "Knowledge management systems" on page 82, a detailed description can be found in the AI literature.
569. See also section 7.3.3 - "Integrating architectures for KMS" on page 311.

- *communication*: as well as coordination and cooperation, e.g., email, workflow management, newsgroups or listserver,
- *document management*: handling of documents throughout their life cycle,
- *access*: to various data sources, e.g., relational data bases, document bases, file servers or Web servers,
- *search:* basic search functionality, e.g., full text search functions for messages, (hypertext) documents, files and folders on file servers or data bases,
- *visualization*: basic functions for a presentation of multimedia files and hypertext documents etc.

A modern, integrated Intranet platform thus can be considered as a KM platform in the sense of a kind of "starter solution" for knowledge sharing. This KM platform comprises the levels Intranet infrastructure including extract, transformation and loading as well as access and security in the KMS architecture presented above[570].

Knowledge management systems in a somewhat narrower sense provide functionality that goes well beyond these basic functions. For each of the following areas, there are a number of application systems or tools respectively which are already available on the market.

Knowledge repositories (knowledge element management systems). Knowledge repositories can be best imagined as document management systems with added features, e.g., with respect to classification and structuring of knowledge elements or with respect to searching with sophisticated filters, user profiles etc. Knowledge is supposed to be embedded in (enhanced) documents and/or forms of discussion data bases. There are three different types of repositories: external knowledge (e.g., competitive intelligence), structural internal knowledge (e.g., research reports, product material) and informal internal knowledge (lessons learned). Knowledge repositories are different from more "traditional" document management systems in terms of added context to information or added functions, such as filtering or synthesizing of the contents. Examples: Fulcrum Knowledge Server (Hummingbird), Livelink (Open Text).

Knowledge discovery and mapping. This category comprises text mining techniques and techniques for representing knowledge sources in a context defined by their relationships reached through clustering, classification and visualization of documents. Example: Intelligent Miner (IBM), Knowledge Miner (USU), Onto-Broker (Ontoprise).

E-learning suites. These systems provide a complete and integrated environment for the administration of tele-learning, both asynchronous and synchronous and to find, catalogue and make available the best expertise within an organization using

570. See section 7.3.3 - "Integrating architectures for KMS" on page 311.

e.g., organizational network analysis, expert yellow pages or expert networks. Examples: E-Learning Suite (Hyperwave), Learning Space (Lotus).

Community builder. These systems help to identify and launch communities or virtual teams independently of the geographical location of team members, provide community homespaces and services. Examples are special tools for moderation, integrated search functionality for distributed messages, contributions to newsgroups and published documents and links in the community homespace. Example: Community Engine (webfair).

Meta-search systems. These systems search different knowledge sources with varying structures (e.g., on-line data base systems, document management systems, file servers, WWW). They offer sophisticated search functions (e.g., the use of meta search data, access to knowledge elements in different systems) and functions for media integration (interface functions bridging differing technologies and formats). Examples: InQuery (Open Text), K2 Enterprise (Verity).

Enterprise knowledge portals. These systems provide access to different knowledge sources and can be best imagined as a "shopping mall" containing a number of "knowledge shops". The portal allows to access these knowledge sources, but does not necessarily integrate all the diverse knowledge sources that can be accessed. Many portals also allow a personalization of the presentation (myportal). Examples are: Hyperwave Information Portal (Hyperwave), Enterprise Information Portal (Hummingbird) or the portal solutions offered by SAP Portals.

Push-oriented systems. These systems contain functions which automatically deliver knowledge elements to participants (e.g., information subscriptions, intelligent agents, support of communities) and thus support the flow of knowledge in an organization. Example: Push Application Server (Backweb).

Collaboration. Expert modeling and decision-making analysis should lead to more collaboration, information expertise and insight sharing among knowledge workers. Systems supporting expert yellow pages and skills directories also fall into this category. Example: Notes (Lotus), Simplify (Tomoye);

Visualization and navigation systems. These systems present relationships between knowledge elements and holders of knowledge. Examples for functions are the presentation of semantic closeness of knowledge elements, the visualization of access statistics to knowledge elements ("beaten tracks"), knowledge maps, mind maps, hyperbolic browsers. Examples for tools are Personal Brain (TheBrain Technologies), Correlate K-Map (Correlate), InXight SmartDiscovery (InXight).

KMS available on the market fall into at least one of these categories. They can be distinguished as well according to their functionality for administration and reporting, e.g., for statistics about the usage of certain functions (e.g., access paths and access statistics to knowledge elements, popular search key words, etc.).

The following exemplary list of KM tools and systems represents a pragmatic *theory-driven classification*. KMS are distinguished according to the main organizational level which they focus on. The list contains a wider set of KM related tools and systems as KMS in a narrow sense span the three levels:

1. **Organization-wide KMS:** enterprise-wide broadcasting systems (e.g., business TV), knowledge repositories, enterprise knowledge portals, directory services, meta-search systems, knowledge push systems (information subscriptions, community support), knowledge visualization systems (knowledge maps), knowledge work process support, e-learning suites, intelligent agents supporting organizational information processing (e.g., for searching organization-external knowledge sources),
2. **Group and community KMS:** community builder and workspaces, ad-hoc workflow management systems, multi-point communication systems (listserver, newsgroups, group video conferencing), collaboration systems, intelligent agents supporting information processing in groups (e.g., in the sense of a transactive memory system),
3. **Individual KMS:** personal search systems (user profiling, search filters), knowledge discovery and mapping, point-to-point communication systems (email, point-to-point video conferencing, instant messaging), intelligent agents supporting personal knowledge management (e.g., for knowledge search).

Last but not least, due to the fact that KMS are developed to support KM initiatives, a typology of KM initiatives can also be used for classifying KMS. Figure B-71 uses the typology of KM focus areas which has been developed by Wiig (1999) on the basis of his extensive consulting experiences and thus a large number of KM case studies. KMS support all KM focus areas even though there is one particular KM focus that relies (almost) exclusively on ICT. A comprehensive KMS certainly can address all focus areas at the same time and thus support KM initiatives of all types. However, each focus area can be supported by a specific bundle of services identified in the architecture for centralized KMS[571].

As in the classification of KM strategies into codification and personalization strategies[572], KMS services are structured according to the respective KM focus areas *ICT resources*, i.e. discovery, publication and integration services, and *person*, i.e. collaboration, learning and personalization services. The KM focus area *process* demands a connection of services according to the needs of the respective business or knowledge process[573]. Finally, the KM focus area *asset* is a mostly strategic and thus the least ICT supported focus area with the infrastructure service

571. See section 7.4 - "Centralized architecture" on page 318.
572. See Hansen et al. 1999.
573. This has been motivated with the concept of knowledge stance in section 6.6.2 - "Activity modeling" on page 250 which helps to conceptualize the assignment of KM services to business processes in a service infrastructure described in section 7.3.2 - "Service infrastructure" on page 304.

asset management as well as reporting and scorecard services that are typically part of business intelligence solutions not considered as KMS in a narrow sense.

asset maximize building and value reallocation of intellectual capital *asset management, reporting, scorecard services*	**process** maximize use of knowledge for operational effectiveness *integration services for business & knowledge processes*
person maximize effectiveness of people-centric learning organization *collaboration, learning, personalization services*	**ICT resources** use IT to maximize capture, transformation, storage, retrieval & development of knowledge *publication, discovery, integration services*

FIGURE B-71. KMS services according to KM focus areas[574]

A classification of KMS can only be considered as preliminary due to the considerable dynamics of the market for KMS. At this stage, the analysis of KMS is a great challenge. This is already visible in the difficulties of defining the term and continues in the trial to present a typical architecture of such systems or to give a comprehensive list of functions. The same is true for a classification of KMS. The pragmatic perspective that KMS are just document management systems with some added functionality which seems to dominate the market is unsatisfying. ICT support for knowledge management is not restricted to the handling of documented knowledge.

Examples for different technological "directions" which provide roots for KMS were suggested in section 7.1 - "Technological roots" on page 273. This list of roots is not complete. It shows from which fields and markets technological support for KM can be expected. Most organizations have installed a large number of application systems and ICT platforms that provide functionality for knowledge management. Especially Intranet platforms form a substantial investment and can provide basic functionality for KM.

Also, many organizations still hesitate to "jump on the bandwagon" as long as it is not clear which KMS vendors will survive the consolidation phase that has just begun and what KMS strategy their main application software suppliers will apply (e.g., Microsoft, Oracle, Peoplesoft, SAP).

574. The KM focus areas used here have been elicited by Wiig (1999, 158) as (1) intellectual asset focus, (2) enterprise effectiveness focus, (3) people focus and (4) information technology and information management focus.

Knowledge management systems might also be viewed as important organizational assets that provide core competencies for the organization. Especially highly knowledge-intensive organizations might view the systematic handling of knowledge in general and their ICT systems supporting KM in particular as their core competence and fear that they might loose a strategic advantage if they implement a standard software solution available on the market.

Most organizations that actually have implemented KMS solutions supposedly have combined several tools and implemented additional functions on their own rather than simply buying specialized KMS software on the market. This leads to the following hypothesis:

Hypothesis 20: The majority of organizations apply organization-specific KMS developments or a combination of organization-specific developments and KMS tools rather than just KMS available on the market.

7.7 Semantic integration

Data and knowledge elements in the data and knowledge source layer typically are scattered across a variety of application systems, e.g., collaboration systems, content management systems, document management systems, file systems and other enterprise systems. Integration of data has been a concern for many years. Relational data base management has unified the way (transactional) data is handled in organizations. The organization of structured, transactional data has been well-understood for years. However, the amount of semi-structured and unstructured data, such as (text) documents, messages, images, media files or Web content has grown substantially and needs to be integrated as well. The integration of these data sources requires other approaches. In addition to data integration, semantic integration provides standards and technologies to integrate knowledge elements from different systems on the conceptual level. Thus, it is not data or Web resources alone that are brokered from system to system, but meta-data about its semantics, its relationships, "meaning" and context. Many of these standards and technologies build on XML.

This section addresses the core integration layer of a KMS[575] that provides access to the heterogeneous data and knowledge sources of an organization in a semantically integrated way, so that knowledge services can be built on top. The integration layer consists on the one hand of function-oriented integration services (function and process integration) and on the other hand of data-oriented integration services (data, user and semantic integration). Data-oriented integration services are the focus of this section. The electronic resources mainly used in knowledge-intensive processes in organizations are semi-structured documents which

575. See sections 7.4.1 - "Overview" on page 319 and 7.4.2 - "Infrastructure and integration services" on page 322.

have to be semantically described using common meta-data standards and semantically rich content and ontology description languages.

Section 7.7.1 discusses the various concepts and elements of the Semantic Web stack. This heads towards machine-understandable semantics and automated reasoning about documents and requires the use of (semantic) meta-data standards for the description of knowledge elements (section 7.7.2) and knowledge modeling, also called the development of ontologies (section 7.7.3). These form the conceptual basis for integration services[576].

7.7.1 Semantic Web

As mentioned above, the differentiation into structured and semi-structured data is often found in e.g., document management or digital asset management (see also sections 2.3.1, 109ff and 4.2.1, 247ff). However, there is no clear demarcation between structured, semi-structured and unstructured data. Generally, all data can be stored in data base systems, even unstructured data in the form of binary large objects (BLOBs). The differentiation is rather of a technical nature. It is postulated that the handling of semi-structured data requires somewhat different technical solutions from relational data base management. These solutions are on the one hand systems specially designed for managing semi-structured data, e.g., content management systems and document management systems. On the other hand, standardization of languages to describe data differ as well. Whereas SQL is the widely accepted standard to define and manipulate structured, data base-oriented data, standards based on XML are used in the realm of semi-structured, content- and document-oriented data.

A number of institutions have developed standards and started initiatives to provide comprehensive frameworks for definition and exchange of meta-data, i.e. semantic information about documents, especially about books, journals, images, photographs, audio and video files. Examples for institutions, standards and initiatives are the World Wide Web (W3C) consortium with XML and the Semantic Web initiative, the International Standardization Organization (ISO) with a large number of standards for document exchange, e.g., the Motion Picture Experts Group (MPEG) 7 meta-data standard for images, audio and video files or the Topic Map standard as well as the Dublin Core standard for exchanging meta-data about text documents which was set up by a consortium including large public libraries.

Structuring, describing, translating, storing and securely accessing semi-structured data as well as reasoning about semi-structured data require a substantial effort. Figure B-72 structures the main technologies that are involved to support these tasks.

Semantic integration of semi-structured data is a complex undertaking. Thus, the Semantic Web initiative breaks down the variety of tasks into a layered structure that helps to understand what concepts have to be defined so that semantic

576. See section 7.4.2 - “Infrastructure and integration services” on page 322.

information about knowledge elements can be easily exchanged between a variety of heterogeneous ICT systems.

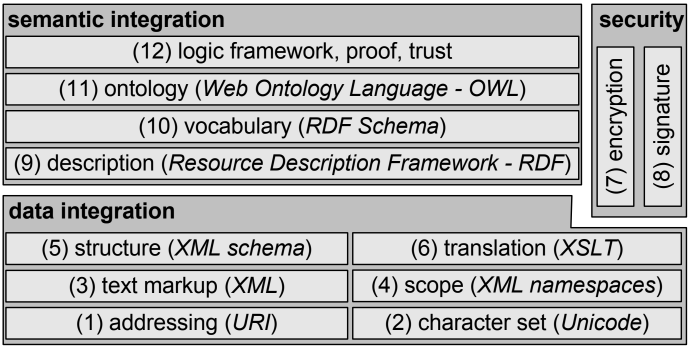

FIGURE B-72. Semantic integration with the Semantic Web stack

Data integration. Data integration requires that agents (users, institutions or applications) exchanging data agree (1) how to address data resources over a network, generally over the Internet, (2) about what character set to use, (3) about the internal structure of documents, called text markup, (4) about the scope or domain in which the specified names in the markup are valid, (5) about how to define a schema, a structure of the elements and attributes in the semi-structured text and (6) how to translate a document that is an instance of one schema so that it conforms to another schema.

In the Semantic Web stack, (1) addressing uses Unified Resource Identifier (URI). URIs are formulated in a standard syntax that is used to uniquely identify objects (or resources) located in any directory on any machine on the Internet, particularly on the World Wide Web, accessed via a specified access method. (2) The Unicode Standard is the universal character encoding scheme for multilingual text. It specifies a numeric value (code point), a name, its case, directionality, and alphabetic properties for each of its characters. Modeled on the ASCII character set, Unicode can encode all characters in all written languages in the world.

(3) The eXtensible Markup Language (XML) is a tag-based markup language for describing tree structures. XML is a set of syntax rules for creating markup languages used to define the structure of documents suitable for automatic processing, i.e. extracting content and structure of XML documents and checking whether an XML document conforms to rules defined by the XML standard, called well-formedness. (4) XML markup defines a vocabulary, also called a markup vocabulary. XML namespaces are a mechanism for creating universally unique names for XML markup vocabularies so that they can be reused by other XML documents. An XML namespace is a collection of names, identified by a URI reference, which is used in XML documents as element and attribute names.

(5) In order to define classes of XML (instance) documents, a number of schema definition languages have been developed that XML documents can be validated

against. The XML Schema language is defined in XML and provides a rich set of data types, extensible by users, for the definition of constraints on XML documents and rules for their construction. (6) The eXtensible Stylesheet Language Transformation (XSLT) provides a standardized way to convert XML documents conforming to one schema, e.g., defined using XML Schema, into XML documents conforming to a different schema. XSLT is also defined in XML and based on the XPath language used to address elements and attributes of XML documents.

Security. Secure access to semi-structured data requires technologies that prevent eavesdropping with the help of (7) encryption and technologies for verifying both, the sender and the receiver of data with the help of (8) electronic signatures.

(7) Encryption uses keys and an encryption algorithm to transform clear (text) data into encrypted data. Data (offline) encryption denotes the permanent codification of data for secure storage, whereas communication (online, wire) encryption is encoding data for secure transfer over networks. Symmetric encryption uses only one key to encrypt and decrypt. Examples for symmetric encryption algorithms are Blowfish, DES (data encryption standard) and AES (advanced encryption standard). Asymmetric encryption uses a pair of a private and a public key. An example for an asymmetric encryption algorithm is RSA (named after its developers Rivest, Shamir, Adleman). XML encryption provides encryption algorithms specifically designed for XML documents, i.e. taking into account the tree structure of XML documents and allowing for data (offline) encryption of sensitive parts of documents as compared to entire documents thus boosting performance[577]. (8) Digital signatures are used to verify the identity of the sender. They are generated using a private key to encrypt checksums of messages. Identity of the sender and integrity of the message can be verified with the corresponding public key. Both mechanisms can be combined to digitally sign the message. XML signature provides a standardized way of signing XML documents[578].

Semantic integration. Based on the standards that support the internal structuring of documents corresponding to a schema, semantic integration aims at providing standards for describing documents or, more generally, Web resources. This is done with the help of (9) statements that describe the resources, (10) a vocabulary for the definition of constraints on these statements, (11) ontologies that show the relationships between the concepts used in descriptions and vocabularies and (12) a logic framework that allows for reasoning about documents and their descriptions.

(9) Based on XML, the Resource Description Framework (RDF) standard is an XML-based language for representing meta-data about Web resources and to relate Web resources to each other. RDF provides mechanisms to add semantics to a resource as a standalone entity without assumptions about its internal structure. Web resources can be text, image, audio or video files, but also things identified on the Web, e.g., items described in a Web page. RDF is based on the idea of identify-

577. See also http://www.w3.org/Encryption/.
578. See also http://www.w3.org/Signature/.

ing things using URIs and describing them with properties and property values. An RDF statement consists of a triple of subject, predicate and object.

(10) The RDF Vocabulary Description Language (RDF Schema) supports designing vocabularies, i.e. classes for instance RDF specifications. Schemas are needed for describing terms used in RDF statements, i.e. types of things, properties and types of things that can be subjects or objects of statements with these properties. RDF Schema proposes well-defined rules for writing these definitions which can be exchanged and parsed automatically to extract semantics of RDF statements about Web resources.

(11) The Web Ontology Language (OWL) is a language for defining and instantiating Web ontologies[579] that include descriptions of classes, properties and their instances. The OWL formal semantics specifies how to derive the ontologies' entailments, i.e. facts not literally present in the ontology. Entailments can be based on multiple distributed ontologies that are combined using defined OWL mechanisms. OWL tools ease the task of applying knowledge representation to building domain ontologies rather than building entire reasoning systems. The normative OWL exchange syntax is RDF/XML, i.e. every OWL document is an RDF document. Compared to RDF Schema, OWL offers more facilities for describing classes and properties, e.g., relations between classes, cardinality, equality, richer typing of properties and characteristics of properties. OWL provides a number of language elements that specifically target the integration of concepts defined in different ontologies. This is especially useful when several systems storing parts of the contents in a KMS have to be brought together. One ontology might be developed per system to capture concept definitions in each system's specific environment. These concepts then might be mapped with OWL, e.g., to provide sophisticated discovery services.

(12) Whereas RDF, RDF Schema and OWL have been standardized for quite some time and there are a number of projects in many organizations that are based on these standards[580], there is still a lot of debate going on at the higher levels of the Semantic Web stack. Standards that allow for specifying entire logic frameworks, exchanging proofs and thus building trust between agents still remain to be seen. Concerning rules, a limited declarative language should standardize the way to query RDF statements. A rule language allows inference rules to be given which allow a machine to infer new assertions from existing ones. A comprehensive logic framework is meant to provide a vocabulary to fully describe and exchange logic assertions over the Web. Additionally, applications or agents can share logic proofs. One agent can send an assertion together with the inference path to that

579. See section 7.7.3 - "Ontology management" on page 387.

580. Examples can be found in Davies et al. 2003, particularly 197ff, or in Tochtermann/Maurer 2006, particularly 249ff, Fensel 2004, 89ff. However, Fensel also sees some shortcomings of OWL compared to other ontology languages, particularly OIL, but predicts that only OWL has a chance of survival (Fensel 2004, 39ff). It should also be noted that there are more ontology languages following other types of logic such as predicate logic, e.g., CycL, KIF or frame-based logic, e.g., Ontolingua, Frame Logic (see Fensel 2004, 21ff and the literature cited there).

assertion starting from assumptions acceptable to the receiver. This requires a standardized language and a standard proof engine. The proof language together with digital signatures signing proofs should turn a web of reason into a web of trust.

7.7.2 Meta-data management

In a simple, yet frequently cited perspective, meta-data are data about data. A KMS contains knowledge elements, i.e. electronic resources of varying types and formats as well as meta-data which give further information about their content and associations. However, one knowledge element's meta-data can simultaneously be another knowledge element's data[581]. There are a number of reasons to assign meta-data to knowledge elements (also Gill et al. 2007):

- *increased accessibility*: Meta-data are a first step to provide meaning about knowledge elements and can be used for smarter information retrieval.
- *retention of context*: The context of a knowledge element is crucial for the reconstruction of knowledge by a user. Knowledge elements can only be correctly interpreted and integrated into a personal knowledge base, if the user can associate the knowledge elements with the correct context.
- *versioning*: Knowledge elements often exist in multiple versions according to storage format and content type, e.g., note of an idea, email message, abstract, research paper, update of the paper, book based on that research paper, learning object, portion of a WBT course[582]. Meta-data help to maintain relations between versions.
- *legal and security issues*: Access privileges and copyright information have to be maintained to assure correct handling of knowledge elements.
- *system improvement and economics*: Meta-data about the usage of knowledge elements can help to improve the system, e.g., by providing shortcuts for often used elements, or to reduce cost, e.g., by automatically transferring little used elements to cheaper storage media.

Meta-data can be used to describe any kind of data from structured to unstructured. The structure itself already is a form of meta-data and usually provides information about the name of the data element, its data type and its relation to other data elements (e.g., an XML Schema for an XML document). Element names are often not sufficient to carry all relevant information. Additional meta-data is needed that either describes the content, e.g., keywords, domain, or the context of

---

581. For a definition of and examples for knowledge elements see section 7.2.1 - "Types of contents" on page 282. The notion of meta-data versus data, though intuitively understandable, has been subject to a lot of attempts for more precise definitions. However, in this section, it seems to be sufficient to distinguish between data and meta-data according to the technical implications that a description of knowledge elements has for semantic integration. For example, the description of a person can be data of her skill profile, but meta-data assigned to her publications. It is often only a technical distinction between both.

582. See also the maturing process of knowledge objects in section 7.2.1 - "Types of contents" on page 282.

the data especially for semi-structured data. The context can be further subdivided into creation context, e.g., author, creation date, project, and the application context, e.g., customer, intended use. Summing up, three types of meta-data can be identified:

- *Content* meta-data relates to what the object contains or is about, and is intrinsic to an information object.
- *Context* meta-data indicates the aspects associated with the object's creation and/or application and is extrinsic to an information object, e.g., who, what, why, where and how aspects[583].
- *Structure* meta-data relates to the formal set of associations within or among individual information objects and can be intrinsic or extrinsic.

The structure is extrinsic in data base tables where data and structure are separated and intrinsic in XML documents where tags and content are mixed. Meta-data can be informal, e.g., free text description, semi-formal, e.g., structured according to a user-defined structure, or formal, e.g., structured and compliant to an organization-wide standard or a standard backed by a consortium of IT companies or a supra-organizational standardization body.

Integration of resources, more specifically knowledge elements, in KMS with the help of meta-data requires a standard language for the serialization of meta-data annotations, a content-oriented standard to define the available meta-data fields and a standard language to formalize an ontology[584]. The latter is used to define the domain and range of meta-data fields and relate meta-data on the type level as well as individual document objects or real-world objects on the instance level by reasoning about the defined concepts[585].

As mentioned in section 7.7.1, many institutions have developed a large variety of meta-data standards. Content-oriented meta-data standards focus on standardization of meta-data fields and can be serialized with the help of languages like XML and RDF[586]. There are a number of domain-independent initiatives to standardize meta-data, e.g., Dublin Core [Hi05], Digital Object Identifier[587], or the Text Encoding Initiative[588]. The Dublin Core Metadata Initiative is an example for a standardization effort primarily aimed at the description of text documents. The standard defines a set of elements that are mainly based on experiences made in public libraries, e.g., Library of Congress, Deutsche Bibliothek. Table B-20 gives some examples for elements in the standard and their descriptions.

Additionally, there are a large number of domain-specific meta-data standards, e.g., in the areas of publishing, library, education, museum or multimedia. Examples are Learning Object Metadata (IEEE 2007), PRISM[589] or MPEG-7[590].

583. See section 7.5.3 - "Example: Infotop" on page 349.
584. For this and the following detailing of these integration tasks see Maier/Peinl 2005.
585. See section 7.7.3 - "Ontology management" on page 387.
586. See section 7.7.1 - "Semantic Web" on page 375.
587. URL: http://www.doi.org.
588. URL: http://www.tei-c.org.
589. URL: http://www.prismstandard.org.

MPEG-7 is a standard that is used to describe multimedia data, especially data stored in MPEG4 video files. The MPEG-7 descriptions of content may include meta-data describing creation and production processes, e.g., director, title, related to usage, e.g., copyright pointers, broadcast schedule, about storage features, e.g., storage format, encoding, on spatial, temporal or spatio-temporal structure, e.g., scene cuts, segmentation in regions, region motion tracking, about technical features, e.g., colors, textures, sound timbres, about the portion of reality or imagination captured, e.g., actors, objects and events, interactions among objects, about how to browse the content in an efficient way, e.g., summaries, variations, spatial and frequency sub-bands, and about interaction of users, e.g., user preferences, usage history. Standards can be compared according to e.g., comprehensiveness, flexibility, languages used for serialization, adoption rate or user friendliness.

TABLE B-20. Examples of Dublin Core meta-data elements

element	description
title	name of the object; could be derived from the filename or from the content
description	abstract or summary of the content in free text form
subject	keywords can be assigned to illustrate topics
creator	entity responsible for authoring the content, e.g., a person, an organization or a service
date	date of an event in the lifecycle of the resource, e.g., creation
relation	links to Web resources (relation.uri) or other sources (relation.other)
language	country code (e.g., us, uk, de) representing the language of the object
rights	e.g., copyright, intellectual property rights, or digital rights (DRMS)
type	categorization, genre or similar aggregation
format	physical or digital manifestation of the object, usually in form of a MIME type

An ontology can be used to relate the meta-data fields. Popular ontology languages include DAML+OIL, Ontolingua and OWL[591]. Ontologies for an organizational KMS can be developed on the basis of existing ontology types, like enterprise ontologies that define organizational structure, domain-task ontologies that define processes, domain ontologies that define relevant topics and common sense ontologies that define e.g., location and time concepts (Gómez-Pérez et al. 2004). Recently, more comprehensive specific ontologies have been proposed for a variety of domains[592].

590. Martínez et al. 2002.
591. See sections 7.7.1 - "Semantic Web" on page 375 and 7.7.3 - "Ontology management" on page 387.

A semantic integration layer in a KMS has to offer services for (1) creating meta-data describing heterogeneous documents, (2) storing it either together with or separated from documents in a repository and (3) retrieving it for inferencing to enable advanced knowledge services. These are discussed in the following (Maier/ Peinl 2005):

Creation. The creation of meta-data in most organizations is primarily accomplished manually. Often, the user is prompted to type in author, title and keywords describing a document before it can be saved to e.g., a DMS. Even more inconvenient is manual creation of an RDF file to annotate e.g., a Web resource. From the perspective of a KMS, a manual approach is not appropriate due to the sheer amount of resources that would have to be annotated. There are some first steps towards (semi-) automated creation of meta-data which either use document-inherent structures and tags like DC-Dot[593] that utilizes HTML tags to generate Dublin Core conforming RDF annotations or sophisticated text mining and language processing algorithms to extract meta-data from content like TextToOnto[594]. There are, however, meta-data fields that can be more easily extracted if the document is structured using an XML format like DocBook[595] that already incorporates most Dublin Core elements.

Storage. Basically, meta-data can be stored either inline, as part of the resource or document that is annotated, like in MS Word or Adobe PDF documents, or document-external, e.g., in a separate RDF file or in a relational data base like in many DMS. XML documents also can store RDF annotations inline using the XML namespace concept. Inline storage is especially advantageous when documents are exchanged between several KMS, e.g., between a company and one of its cooperation partners or in a peer-to-peer scenario where resources are stored in a distributed environment. In this case, the sending KMS packs all resource descriptions relevant for the target environment together with every exchanged knowledge element which can then be extracted by the receiving KMS.

This could also be called a process transferring explicit knowledge between different contexts. This is a specific case of more general business processes and, in a more detailed perspective, workflows between organizational units or even between organizations. Theoretically, this is the traditional realm of workflow management. Historically, there have been three distinct perspectives on workflows implemented in workflow management systems (Jablonski et al. 1997, 91f):

592. Examples which are of special interest for the integration layer of KMS are publication descriptions using BibTeX in OWL (e.g., http://zeitkunst.org/portfolio/programming/bibtex_owl/) or the AKT Portal ontology that describes academic researchers, their publications and projects (http://www.aktors.org/ontology/).
593. http://www.ukoln.ac.uk/metadata/dcdot/.
594. http://sourceforge.net/projects/texttoonto.
595. http://www.docbook.org.

- *structured sequences of steps* which sets the focus on events, conditions and actions in the control flow, on disaggregation of tasks, their delegation to agents or resources and particularly on coordination of the tasks,
- *actualized conversation types* which is based on speech-act theories[596] and views workflows as regulated and structured exchanging of speech acts between agents and thus focuses agents and their possibilities to send messages and to react to messages,
- *migrating or stateful objects* which is also called object migration or information sharing model and focuses the objects, documents and their states that determine the actions that can, should or must be performed on the objects and is best to be used in cases where the workflow is determined largely by the type of document.

Inline storage of meta-data for knowledge transfer can be seen as the first step in "activating" resources in general and electronic documents in particular that has been subject to a number of research approaches. Basic idea is to store explicit knowledge about resources, their contexts, as well as application logic directly and inseparably with the resources which leads to the concept of smart, intelligent, living or active document[597]. Active documents represent another variant of the perspective on workflows as migrating objects. The difference to traditional workflow management is that active documents do not require central coordination, but carry all meta-data that is relevant for the workflow with them. These approaches differ with respect to the degree of activation as depicted in Figure B-73[598]:

- *Passive documents* are containers of data with no capabilities to influence the control flow of a receiving system. Meta-data annotations can only be stored separated from the document, e.g., implicitly in a folder structure or explicitly in a relational DBMS.
- *Enriched documents* contain their meta-data annotations as part of the document. With increasing adoption of XML-based document formats[599], this is the primary standard used in order to structure documents.
- *Reactive documents* can trigger (simple) actions for corresponding pre-defined events, e.g., adaptations to user environments, such as language, font size or image resolution.
- *Active documents* contain data, meta-data and application logic or are directly connected to application logic which is a fixed part of the document, cannot be separated from it without substantial loss of meaning, is transferred with the documents and can actively trigger, control or execute functions[600].

---

596. For speech-act theory see Austin 1962, Searle 1969, adapted to electronic communication and computers by Winograd/Flores 1986.
597. Carr et al. 2003, Schimkat 2003, Maier/Trögl 2006.
598. See Schimkat 2003, 54ff, Maier/Trögl 2006, 6ff.
599. Examples can be found in standard office systems, e.g., Adobe Intelligent Document Platform, Microsoft Office or OpenOffice.

- *Proactive documents* additionally attempt to autonomously achieve pre-defined goals. Actions to achieve the goals are integrated into the document. The sequence of actions is selected autonomously according to the reactions of the respective environment[601].

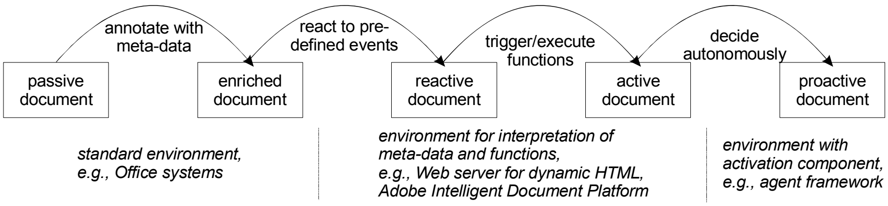

FIGURE B-73. Degrees of activation of electronic documents[602]

The approaches for the implementation of active documents differ not only with respect to the degree of activation. Primary goals vary from simple storage and retrieval of documents over adaptation of the presentation of documents to control of task and workflows. Theoretically, the approaches are all based on meta-data concepts and some aspects of object-orientation, but differ in their reliance on network theory, service-oriented architecture or agent theory. The technical environments for execution of active documents vary on the server side from standard Web servers over vendor-specific servers, especially document management systems, to specific middleware. Exceptions are solutions for proactive documents which rely on an agent platform. On the client side, the approaches rely on widely available platform-independent application software, e.g., Web browser or Adobe Acrobat Reader or on vendor-specific office systems. Documents are mostly realized as containers with diverse formats whereas meta-data are consistently represented in XML-based formats. Not surprisingly, the approaches differ mostly with respect to how they implement functions. Solutions include diverse macro, script and programming languages, the Web service concept as well as software agents.

However interesting and promising approaches for activation of documents are with respect to integration of several document bases and transferring explicit knowledge to other contexts, it is not efficient to store meta-data only inline or in separate files for searching large document collections. Thus, the need arises for a way to store meta-data, e.g., RDF triples, in and retrieve it from a data base. In general, either relational, object-oriented, XML-based data bases or proprietary data base formats can be used. In a KMS setting, relational data bases might be pre-

600. Examples for technical implementations of active documents are Web servers interpreting dynamic HTML documents, adaptive hypermedia systems (Brusilovsky 1996, DeBra et al. 1999, Brusilovsky 2001) or the Placeless Documents approach (Dourish et al. 1999, Dourish et al. 2000).
601. An example is Schimkat's research prototype for living documents, a middleware which is based on the agent framework Okeanos (Schimkat 2003).
602. After: Maier/Trögl 2006, 8.

ferred due to their dominance and the fact that common drawbacks for XML storage like missing white space preservation or breaking digitally signed contents do not seem to be an issue here. Thus, this approach is examined closer (Melnik 2001):

One method would be to store all RDF triples in one table which results in denormalized data. Separate tables for resources, literals, namespaces and statements would dramatically decrease required storage capacity, but also decrease performance as a number of computation-intensive joins have to be made. The Jena toolkit uses the former approach, whereas Sesame is an example for a tool that implements the latter approach. Finally, one could also store RDF data in a data base schema according to the RDF schema describing the structure of the RDF file. This potentially results in a large number of tables and makes it more difficult to retrieve statements independently from their RDF schema, but can also improve retrieval for a fixed and small number of schemas.

Retrieval. Established query languages like SQL, OQL or XPath/XQuery could be used in order to retrieve meta-data from the data base, depending on the type of data base management system used. However, there are many shortcomings that could be overcome with a query language that explicitly supports the RDF triple structure and other RDF language constructs. A number of proposals for such languages have been made, e.g., iTQL, RDFQL, RDQL, RQL, SeRQL, and SPARQL. Although these languages look similar, since they all imitate SQL, their capabilities are quite different.

Haase et al. (2004) evaluate a number of these languages and define the following requirements for an RDF query language: support for (1) RDF abstract data model, (2) formal semantics and inference, (3) XML schema data types for literals and (4) statements about resources. They further judge the languages according to their (5) expressiveness, (6) closure, (7) adequacy, (8) orthogonality and (9) safety. They conclude that especially grouping and aggregation, as well as sorting and optional matching are poorly or not at all supported. Also, RDF language elements like XML data types, containers and reification are only supported in a few cases. From a KMS perspective, language capabilities and industry support are important criteria. Stier's (2005) evaluations supervised by the author[603] as well as the updated results of Haase's (2005) research show that RDFQL scores better than other query languages. Nevertheless, it seems that either RDQL, due to its support by HP and implementation in several tools, e.g., Jena, RDFStore, Sesame, 3Store, RAP, or SPARQL due to its progress in the W3C standardization process will become widely accepted.

There are a number of tools available that support RDF storage and retrieval, most of which are the results of academic research and are freely available (Stier 2005). However, maybe as a result of that, only few tools are easy to use, most of them even lack a graphical user interface. Some remarkable exceptions are 4 Suite,

603. See also Maier/Peinl 2005.

Sesame, KAON and Kowari. Sesame supports RDQL as well as RQL and SeRQL, whereas most other tools only support one language. This is especially interesting for KMS for flexibility reasons, as long as there is no clear standard yet. RDFQL support is only available in the commercial tool RDF Gateway from Intellidimension. A prototype implementation developed at the author's department builds on top of Jena and enhances the toolkit with a Web-based client for retrieval as well as a Java-based graphical client with support for creating, storing and retrieving RDF from the data base (Stier 2005, Maier/Peinl 2005).

Figure B-74 summarizes the most important steps for creation, i.e. annotation and serialization, storage and retrieval of meta-data from the perspective of the integration layer in a KMS.

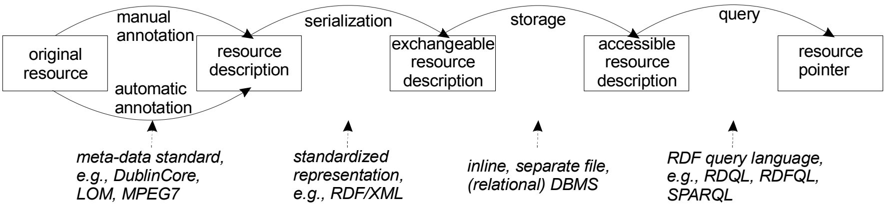

FIGURE B-74. Meta-data management

The integration layer in a KMS builds on semantic descriptions of documents to provide functionality to the knowledge services on the upper layers, such as semantically relating knowledge elements to each other or identifying experts based on authorship. Thus, creation, storage, retrieval and processing of meta-data and associated ontologies are required. With XML, RDF, OWL and RDF query languages as well as the use of content-oriented meta-data standards, a significant part of the required integration services can be realized.

However, the lacking standardization of RDF query languages together with missing capabilities of the proposed standards and insufficient tool support inhibits a broad implementation of semantic integration layers in organizations. Moreover, despite a number of content-oriented meta-data standards that seem well-suited for their designated domains, there is no broadly accepted standard that covers all relevant aspects in a KMS context.

The various developments in the technology-enhanced learning community on the standardization of learning objects, learners etc. as well as modularization that is already designed for some standards seem to be important steps in this direction. Probably a more flexible, modular meta-data annotation system with a few basic attributes for all documents together with a set of document type-specific attributes could link standards for specific domains. The meta-data should be organized in a kind of top-level ontology according to the identified categories and the dimensions time, topic, location, person, process and type[604].

This already points towards the concept of ontology management. On top of meta-data management which governs the basic services of semantic integration, ontology management can provide for semantically richer descriptions of resources, their properties, relationships and rules that allow for reasoning among the knowledge about resources.

7.7.3 Ontology management

The concept of ontology has already been defined and its impact on KM has been discussed in section 6.6.3 - “Knowledge modeling” on page 257. In this section, the model of an architecture for ontology-based knowledge management systems and a procedure model for ontology development are briefly introduced to illustrate how these concepts can be implemented in a KMS.

Figure B-75 shows a procedure model for developing ontology-based KMS which is based on several related procedure models (Sure/Studer 2003, 42ff) and has been applied in several projects at the AIFB institute, University of Karlsruhe (Germany). It gives guidance in developing a core knowledge structure that forms the basis for semantic integration of numerous data and knowledge sources.

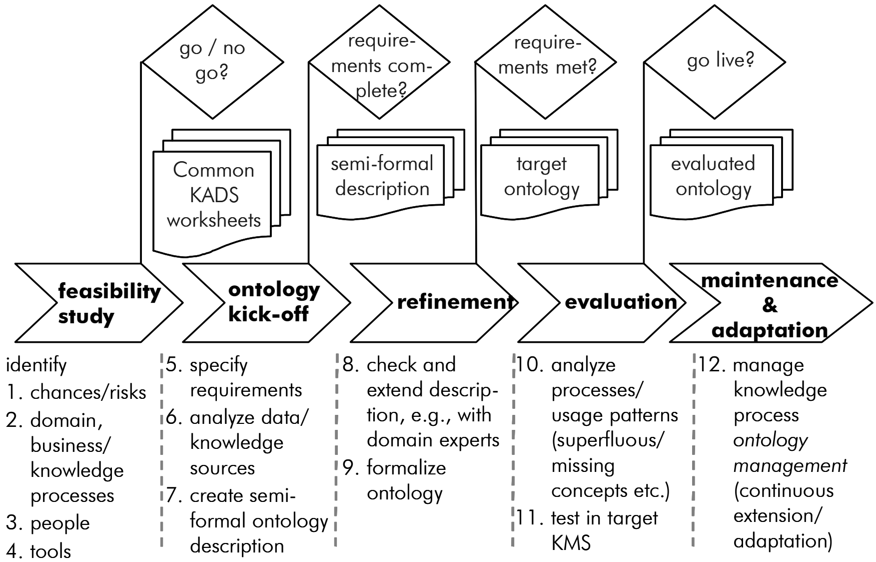

FIGURE B-75. Procedure model for ontology development[605]

Feasibility study. This process step in the procedure model aims at setting the scene for the ontology development project and leads to a decision about continuing the effort or terminating the project. Setting the scene means to determine scope and domain of the project as well as the people involved. Domain experts

604. See section 7.5.3 - “Example: Infotop” on page 349 and Maier/Sametinger 2007.
605. After: Staab 2002, 204, Sure/Studer 2003, 34.

have to be selected from the group of the latter. Also, the ontology language and the tools to be used have to be selected. Criteria for selecting an ontology language are e.g., (1) intuitiveness for domain experts and other users that participate in the modelling effort, (2) existence of a well-defined formal semantics with established reasoning properties in terms of completeness, correctness and efficiency and, increasingly, (3) the possibility to serialize in XML, RDF and OWL (Fensel 2004, 48). The main tool for ontology construction is an ontology editor during build-time, for example Protégé or OntoEdit, whereas a reasoning system, e.g., OntoBroker[606], is needed during runtime for the provision of inference services based on the ontology and thus answering questions with the help of the ontology.

Ontology kick-off. In the kick-off phase, requirements are specified that the ontology has to fulfill. Data and knowledge sources which should be integrated with the help of the ontology are identified and studied. Finally, a semi-formal ontology description is created that considers the requirements and covers all data and knowledge sources deemed relevant.

Refinement. The refinement phase relies heavily on support by domain experts and by ontology tools. First, the semi-formal ontology description is checked for consistency, completeness, relevance and other criteria. In the following step, a decision has to be taken concerning which ontology languages to use. In the meantime, most projects use some form of standard ontology language, e.g., OWL, as the basis. However, in many cases, additional ontology elements are used which are not covered by the standards. In this phase, also the support by an ontology editor and, subsequently, the existence of a corresponding ontology inference engine, are important prerequisites for an economically feasible ontology management process. The refinement phase concludes with a check whether the formalized ontology fulfills the requirements stated in the kick-off phase.

Evaluation. In the evaluation phase, the ontology is tested with the help of test cases simulating typical queries, certain usage patterns and processes. This should allow for a check whether all concepts are needed and/or whether additional concepts are necessary in order to support the patterns and processes deemed necessary. If revisions are required, then a switchback to the refinement phase might be necessary. Finally, the ontology is deployed in the integration layer of the target KMS in order to check whether it fits the environment in which it should be used.

Maintenance and adaptation. The final phase revolves around an operational system using the ontology. As ontology management is quite resource-consuming, it is crucial that maintenance and adaptation are planned systematically and supported by methods and tools. Whereas the ontology management process so far has the character of a project, it is this phase that points towards continuous management with respect to the integration layer of a KMS. However, there are less expe-

606. http://www.smi.stanford.edu/projects/protege/, http://www.ontoprise.de/

riences with this phase as compared to the earlier phases (Staab 2002, 205). Larger adaptations require another design cycle of refinement and evaluation phases.

Once the ontology is developed, it is deployed in a KMS. Figure B-76 shows components of a KMS from the perspective of ontology management.

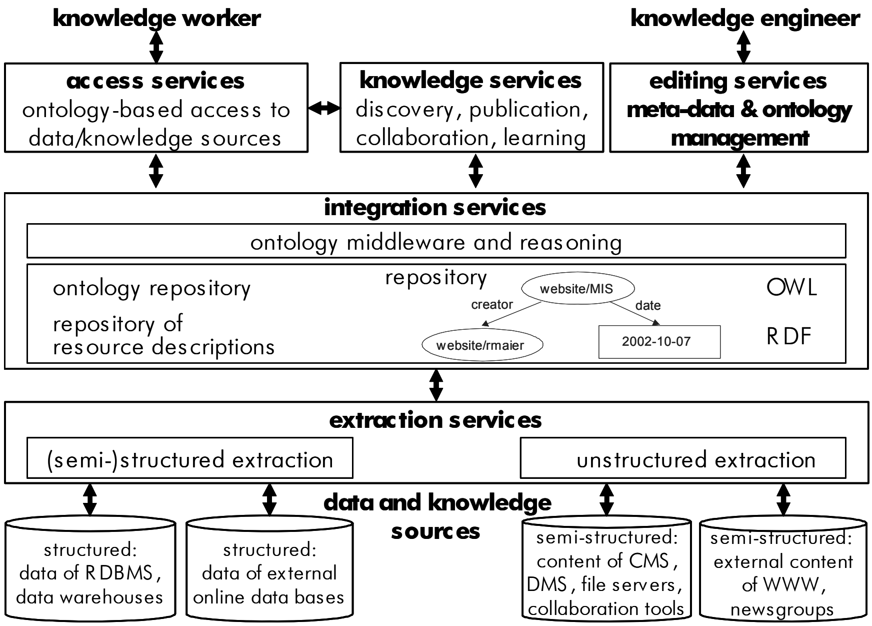

FIGURE B-76. Architecture for ontology-based KMS[607]

Extraction services are required in order to bring knowledge elements from structured and unstructured sources of data and knowledge or, more precisely, descriptions thereof and references to these elements into the repository of resource descriptions. This repository that also contains the integration ontology is the basis of integration services which are implemented with the help of an ontology middleware and reasoning systems. This is the central layer that realizes the inferencing services required by knowledge services built on top of them. Figure B-76 also shows the two main roles of a knowledge worker and a knowledge engineer which either use the system for an improved access to data and knowledge sources or for meta-data and ontology management processes.

607. This architecture integrates the architecture for Semantic-Web-based KM after Davies et al. (2003a, 6) with the architecture presented in Figure B-59, “Architecture for centralized KMS,” on page 319.

7.8 Résumé

This chapter was dedicated to the analysis of ICT tools and systems that can be used in a KM initiative. First, the *technological roots* of KMS were analyzed. The roots of KMS can be found for example in business intelligence, document or Web content management, communication, Groupware, learning management, portal, search and retrieval, visualization, workflow management and—last, but not least—artificial intelligence and knowledge-based technologies. Central to each KMS implementation are the *contents* that are managed by these systems. A list of sixteen types of contents will be used in the empirical study:

- knowledge about organization and processes,
- product knowledge,
- patents held by the organization,
- external patents,
- internal studies,
- external studies/reports,
- lessons learned,
- best practices,
- ideas, proposals,
- questions, answers (FAQ),
- employee yellow pages/skills directories,
- knowledge about business partners,
- directory of communities,
- internal communication,
- external on-line journals,
- private contents.

This list is not complete. There are also many more ways to classify contents some of which were addressed in this chapter[608]. The definition of the smallest unit of explicit, documented knowledge, the *knowledge element*, was found to be challenging. Unfortunately, we are still far from an agreed upon understanding of what exactly is stored and handled in a KMS both, in the literature and in the organizations. A number of examples for knowledge elements were listed that will be used in the empirical study, such as a document containing lessons learned, patents, a description of skills or of a best practice, a contribution to a newsgroup, an element in an experience data base or an entry in a list of frequently asked questions and answers. The description of a knowledge maturing process can help organizations to analyze their KMS, define types of knowledge elements on different levels of maturity and systematically manage maturity paths between them.

The size of the contents of a KMS will be measured in terms of the number of knowledge elements handled and in terms of the storage capacity used. Contents

608. See also section 4.2.2 - "Types and classes of knowledge" on page 66.

can be stored using different *types of media*, such as documents, multimedia contents, contributions to newsgroups or data base elements.

The structuring of an organizational knowledge base is often considered one of the key tasks in KM. However, as the current state of theory does not provide easy-to-use methods and instruments to aid this task, the investigation into the state of practice of knowledge organization and structuring will have to be limited. The number of knowledge clusters will be related to the number of knowledge elements. Additionally, organizations will be asked what ways of structuring the organizational knowledge base—hierarchical or network—they apply.

The concepts of KM service and a KM service infrastructure have been introduced and link the design of KMS with KM strategy, business and knowledge processes. Due to the substantial interest in KM and the subsequent vagueness with which the term KMS is used especially by vendors of ICT tools, platforms and systems to support KM, it is not surprising that there are a number of architectures proposed for KMS. An ideal *architecture* for a centralized KMS was presented that integrated theory-driven and market driven architectures as proposed in the literature and vendor-specific architectures that have been developed with one particular KMS in mind. The architecture consists of five layers that build upon each other and reflect the substantial complexity of KMS solutions in practice: (1) access services, (2) personalization services, (3) knowledge services, (4) integration services and (5) infrastructure services that build on data and knowledge sources. A comprehensive list of functions of KMS was presented, structured according to the architecture's layers. The architecture together with the list of functions can be used as a checklist to evaluate KM tools and systems. Due to the importance of the integration layer, a separate section deals with the Semantic Web initiative, metadata and ontology management as the main pillars of semantic integration in KMS.

Evaluation of the quality of contents and functions of a KMS can be supported by structured lists of criteria for information quality. Table B-21 assigns the criteria for information quality to the five layers distinguished in the architecture for centralized KMS.

TABLE B-21. Assignment of quality criteria to levels of KMS architecture

level of KMS architecture	**information quality criteria**
access services	accessibility, security
personalization services	applicability, conciseness, convenience, timeliness
knowledge services	accuracy, clarity, currency, interactivity, traceability
integration services	comprehensiveness, consistency, correctness
infrastructure services	maintainability, security, speed

Infrastructure services are evaluated according to their contribution to maintainability, security and speed of knowledge "transmission". Integration services provide comprehensive, consistent and correct knowledge. Knowledge services improve the knowledge's accuracy, clarity, currency, interactivity and traceability. Personalization services foster the applicability, conciseness and convenience of the knowledge presented to the knowledge worker. Also, push services provide timely knowledge. Access services obviously primarily deal with the accessibility and security of knowledge.

The ideal architecture of a centralized KMS was contrasted with an architecture of a distributed or peer-to-peer KMS. The proposed advantages concerning acceptance, flexibility and cost can only be realized if the substantial problems of a decentralized management of meta-data and the lack of semantic integration of the knowledge elements in this architecture can be overcome. Still, this is a promising direction for future research that might remove the barriers to use a (costly) centralized KMS solution, especially by small and medium-sized enterprises and for collaboration and knowledge sharing across organizational boundaries.

Functions of KMS can be categorized in a multitude of ways. In the following, an example for a pragmatic classification will be discussed which will be used in the empirical study. Once again, the differentiation between a technology-oriented and a human-oriented KM approach is visible in the distinction between groups of *integrative* and *interactive KMS functions*. In addition, there is one group that bridges these two groups of functions. There is also a group of KMS functions which can easily be classified as integrative, interactive or bridging functions. The links to KM tasks and processes are shown for every group[609].

Knowledge integration. These functions support knowledge processes. Examples are knowledge publication, structuring and linking, contextualization, quality assurance, storing and feedback:

- *knowledge search and presentation*: keyword search, meta-search system, user-initiated filters, navigation, information subscriptions for interested users, thesaurus/synonyms, presentation of new/unread documents, search assistants / search support, three-dimensional visualization, semantic closeness between knowledge elements, ranking of knowledge elements, presentation of full texts,
- *knowledge acquisition, publication and organization*: publication of pre-structured contents by participants, publication of not pre-structured contents by participants, indexing/integration of published contents, comments to knowledge elements, manual import of external knowledge elements, automatic import of knowledge elements from external sources, generation of knowledge elements from internal data sources, statistical data analysis, knowledge repository, automatic indexing of full texts, automatic classification/linking of knowledge elements, semantic analysis of knowledge elements, (hyper-) linking of published contents, structuring and management of knowledge clusters,
- *CBT*: computer based training.

609. See section 6.3 - "Process organization" on page 207.

Knowledge interaction. These functions support knowledge processes and knowledge-intensive business processes. Examples are asynchronous and synchronous communication and cooperation, person-to-person, team, community and enterprise-wide communication and cooperation, expert brokering:

- *knowledge communication and cooperation*: email, email distribution lists, listserver, ad-hoc workflow management system, newsgroups, point-to-point video conference, multi-point video conference, networked group video conference rooms, audio conference, group conference management, list of participants currently on-line, instant messaging, chat, electronic whiteboard, application sharing, co-authoring functions, electronic brainstorming,
- *tele-learning*: videoserver, live broadcasting of videos.

Functions bridging knowledge integration and knowledge interaction. This group of functions supports knowledge processes and knowledge-intensive business processes. They attempt to close the gap between integrative and interactive KMS by e.g., supporting direct interaction between participants and e.g., authors of knowledge elements, by using other participants' access patterns to integrative KMS or by integrating knowledge structures and knowledge networks into comprehensive knowledge maps. One particular sub-group of functions bridging knowledge integration and knowledge interaction can be called knowledge profiling: system-initiated automatic participant-oriented selection, repackaging and presentation of knowledge elements, push-technologies, automatic and participant-initiated building of user, group, team or community profiles, topic-oriented information subscriptions and the like:

- *knowledge search and presentation*: intelligent agents, user profiles, access statistics for knowledge elements, presentation of knowledge elements in maps, access paths to knowledge elements/clusters, presentation of related knowledge elements, navigation from knowledge elements to authors/communities,
- *knowledge acquisition, publication and organization*: feedback from participants to authors, automatic notification of potentially interested, definition of roles for participants,
- *administration*: role-specific configurations of KMS.

Knowledge management. These functions support the knowledge management process. Examples are identification and visualization of enterprise-wide knowledge, reporting of the use of the infrastructure, identification of knowledge gaps, enterprise-wide knowledge quality management (e.g., definition of criteria to be met for publishing knowledge elements) or administration of KMS.

This rather high-level classification can be detailed at will. However, interviews with knowledge managers suggest that to date organizations use only a couple of functions to support the administration of KMS, reporting and visualization (in the form of knowledge maps). The more advanced functions are supported at most with the help of prototypes. Moreover, these reporting, administration and visualization functions can be assigned to the three categories integrative, interactive and bridging KMS functions:

- *knowledge integration:* reports concerning knowledge elements,
- *knowledge interaction:* reports concerning participants or collectives of participants,
- *functions bridging knowledge integration and interaction:* administration of group, team and community profiles and privileges; personalization of user interface; development and management of knowledge maps.

Also, a number of approaches to the *classification* of KMS or, more generally, to the classification of ICT tools supporting a KM initiative were reviewed. The classifications comprised more traditional ICT, ICT platforms, specialized KM tools as well as comprehensive, integrated KMS solutions. Then, an amalgamated classification on the basis of the literature review as well as the market study on KMS was suggested (see Maier/Klosa 1999c). This classification leaves more traditional ICT tools, systems and platforms out of consideration[610]. It distinguishes knowledge repositories, knowledge discovery and mapping, e-learning suites, community builder, meta-search systems, enterprise knowledge portals, push-oriented systems, collaboration as well as visualization and navigation systems. Additionally, a theory-driven broad classification divides ICT to support KM on the organizational level, the group and community level as well as the individual level. Finally, an empirically motivated classification classifies ICT according to Wiig's KM focus areas.

Due to the considerable dynamics of the market for KM-related ICT, it must be noted that all these classifications can only be considered as preliminary. Several vendors currently attempt to integrate as many KMS functions into their KMS as possible[611]. These integrated KM platforms or KM suites bridge classes and combine e.g., a knowledge repository, an e-learning suite, a meta-search system, an enterprise knowledge portal, a push-oriented system as well as a visualization and a navigation system. It seems that architectures, lists of KMS functions and classifications presented in this chapter together provide a good foundation to structure the market of KM-related ICT.

610. The more traditional ICT supporting a KM initiative was discussed in section 7.1 - "Technological roots" on page 273.
611. Recently, many vendors, e.g., Open Text, have acquired or merged with a number of other vendors in the KMS market to speed this integration process.