Enterprise Application Integration: New Solutions for a Solved Problem or a Challenging Research Field?

Joachim Schelp and Frederic Rowohl

University of St. Gallen, Institute for Information Management, Switzerland

Tel: +4 (171) 224-3350, Fax: +4 (171) 224-2189, {Joachim.Schelp, Frederic.Rowohl}@unisg.ch

ABSTRACT

When closing the loop back to legacy systems data warehousing is becoming a general IT integration topic, which is—partly—discussed under the label enterprise application integration. This article enables the interested reader to identify current problems in enterprise application integration. It shows solutions reached in previous research efforts as well as solutions provided by today's software vendors. Finally, it sums up the pending gaps and open research fields.

INTRODUCTION

For many years, the integration aspect has been a main topic in both theory and practice of computer science. The overall aim is to integrate heterogeneous applications across processes in the whole enterprise. During the last decade, several approaches have been made in both research and practice to establish such integrated information systems [Mert66, Sche95, Mert95 etc.]. The theoretic approach often is to propose new application systems or architectures. E.g. *Scheer* [Sche95] and *Mertens* [Mert95] worked on reference models for the integration of applications (and data). Other approaches focus on data flows, processes or functions.

Still today, there is an increasing demand for integrating applications in the business area. Pending integration efforts like data warehousing projects are to be extended to bring their results back to legacy systems, e.g. to feed customer relationship management systems. New business models require the adoption of new technologies without leaving enough time to build up completely new systems: existing applications have to be extended by connecting them to new application types bearing these new technologies. The development of horizontal applications is an example of this trend [Wint00].

Although several approaches were developed to integrate existing applications, software vendors still present new toolkits based on new technologies to address the specific needs of enterprise application integration problems. As these technologies are not mentioned explicitly in the existing approaches for application integration, we have to ask whether these approaches fulfill today's needs and represent the state of the art in application integration. Main problems occur in both increasing complexity and dynamics of today's businesses and information technologies as well as in varying decentralized architectures. Furthermore, for most of the approaches presented in research, existing applications or architectures are discarded in favor of new ones without referring to or even *integrating* the existing set of systems.

This article derives specific problems occurring from integration and requirements for the integration of existing heterogeneous applications regarding current business needs. After an initial discussion we will show the result reached in several research projects in the past and compare them with the problems and requirements described before. In a third step we ask for current solutions provided by both software vendors and the scientific community. Again we will compare these solutions with the problems and requirements described initially. Finally we show the gaps still pending and deduce the open research problems. A short conclusion will round up this paper.

CURRENT REQUIREMENTS FOR INTEGRATION OF APPLICATIONS

As stated above, still today integration is an urgent need for companies. This has several reasons which result from the rapidly evolving technological environment and the fast adoption in business processes:

- New technologies enable new application types opening new distribution channels for companies. E.g. in the banking area there are not only ATM-networks as an additional front-end to the customers but also WWW- or WAP-interfaces. And again new technologies like UMTS are at the horizon [EITO01].
- Because of the internet there is an urgent need for enabling online transactions instead of the old-fashioned batch processes. Batch processing methods require time frames which customers do not accept when initiating online transactions.
- New applications for customer relationship management extend the need for transferring analytical data from data warehouses back into transactional systems. This remarkably impacts performance, access, and data organization for these systems. Data has to be organized by subjects and not only by relations, online transactions are required as well as short request times. These are challenges for both analytical and transactional systems, which today fulfill only one or two of these requirements.
- Integrated solutions like enterprise resource planning systems reach their limits when new technologies must be incorporated. Therefore additional tiers must be introduced in the information system environments to add an additional transformation layer between old and new applications.

To sum up these problems, we can derive the following requirements. The above problems can be united to *complexity* (in different ways), *compatibility* (e.g. to exchange data), and limited *capacity* of existing systems.

- The higher legacy systems' data, function, and process models complexity is, the more complex is an integrated logical view on them. But this is required to couple the systems together. Therefore there is a strong requirement to build integrated logical models of different levels of complexity.
- Especially old legacy systems have limited means of importing or exporting data in foreign formats. To exchange and convert data between different systems metadata has to be collected, structured, and exchanged. Therefore methods and tools for metadata management are required which can enable data quality issues as well.
- From a more technical point of view additional application layers are required as the older legacy systems are often working at their capacity limit. Because of little system performance, sometimes

they cannot export and import data to additional systems even if they have the functionality to do this. The resulting enhanced IT landscape of a company must be planned and maintained.

These requirements are a challenge. As technological improvements are still at the beginning, companies are urged to work with continuously changing environments. Heterogeneity in the IT-landscape within an enterprise will be an ongoing challenge which has to be met by IT-concepts and strategies as well as by middleware applications in use.

The addressed requirements can be met on a more conceptual and on a more technical level. Conceptual approaches are discussed in the following section, technical ones in section 4.

EAI IN RESEARCH AND LITERATURE

Since the early 70s the need for integration and concepts of integration of enterprise applications is a broadly discussed topic in the computer science community. Several existing suggestions have been made and further developed in literature. However, not all were implemented in practice, so a proof of concept is not been made for each proposed solution. With today's increasing demand for integration models the view of existing approaches could give an indication of recent research efforts. This chapter gives an overview of existing integration models and procedures.

Different views on integrated information management can be distinguished [Mert95]:

1. *Integration object.* The integration object can be data, functions, processes, methods, and programs.
2. *Direction of integration.* The direction of the integration can be horizontally or vertically.
3. *Range of integration.* The integration can take place in one area of an enterprise, in the whole enterprise, or between different enterprises.
4. *Degree of automation.* The intended degree of automation can be full automation or automation in parts.

Concerning the integration object, recent models focus on both data and functions, and partially on ideas. The other views are secondary and often follow the integration object. Different levels of integration of data and functions are shown in Figure 1. In the left figure two systems exchange and process data automatically. This is a very simple way of integration, but the data has to be stored redundantly. In the middle different applications use the same database. The data has to be modeled in an appropriate way to integrate it in one database. In the right figure the functions are linked together on a technical level. The result is an integration of functions (see also [Mert95, p. 24]).

Basically, existing integration models can be divided in data oriented, function oriented, and others [MeHo92].

Data Oriented Integration Models

1. Conceptual Data Model" by *Vetter* [Vett90]. *Vetter* separates in his Conceptual Data Model different phases of development, namely the design of the object system, the information system, the database concept, and the processes. He defines a bottom-up approach when he develops the enterprise wide data model by joining the separated data models of applications from the business departments.

Figure 1: Integration of data and functions (Mert95)

two modules, double data   two modules, single data   one module
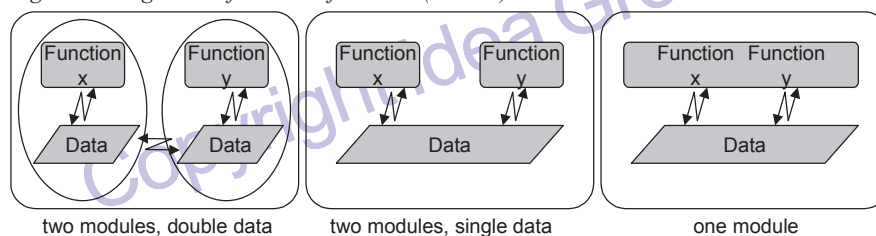
The aim is to develop an extensive data model of the informational environment.

2. „ARIS" by *Scheer* [Sche91]. *Scheer* developed with his „Architecture of Integrated Information Systems (ARIS)" a framework for development, optimization and realization of integrated applications. The „ARIS-House" (Figure 2) is an architecture combining multiple views and levels of integration.

Figure 2: The "ARIS-House" (Sche91)
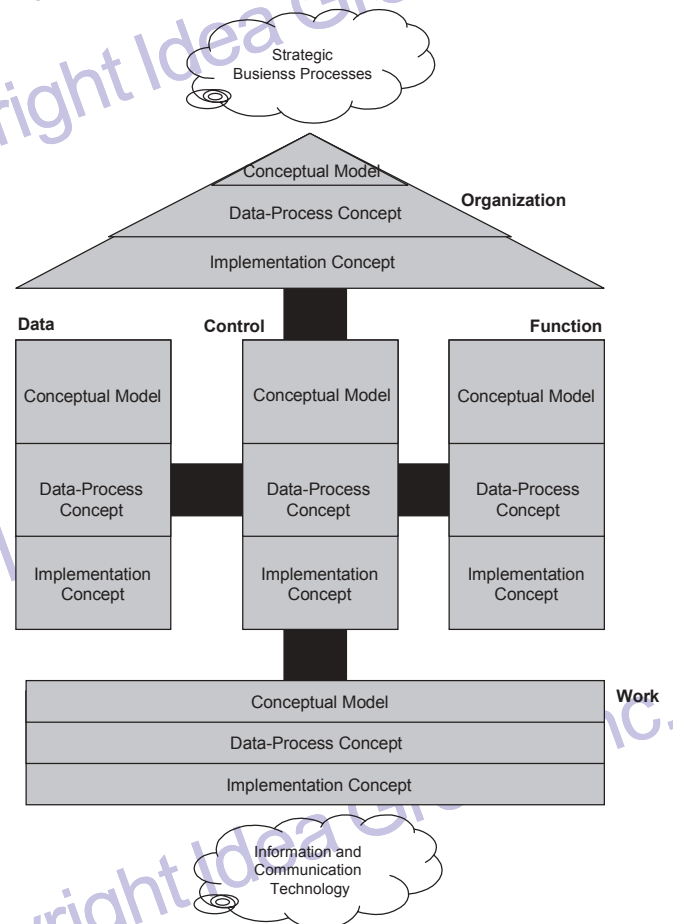
He defines process chains as starting point of the development of the system architecture. The process chains are assigned to separated levels: on the application level there are single proceedings, on a higher level the process types and the standard processes. On this level the enterprise information systems are allocated. On a meta level the general process model is defined.

The development of ARIS consists of four phases. First is the analysis of the process chains. Second is the modeling of the conceptual model, followed by the design of the data-process concepts. Fourth and last step is the design of the implementation concept. The phases 2 to 4 are subdivided by different views (function, organization, data, control). The conceptual data model is an important element of ARIS, whereon the focus on data integration is based.

ARIS consists of entity relationship models for different functional areas, in detail for production, technology, procurement, sales,

human resources, accountancy and administration. The ARIS-approach is top-down because of the definition of process chains from an external point of view.

3. UFM-UDM by *Rauh* [Rauh90]. *Rauh* developed a function model of the enterprise (UFM) and a data model of the enterprise (UDM) and combined them. The UDM should be a complete description of the enterprise from a data point of view. The data are assigned to functions as procurement, production, sales, human resources and accountancy, management, controlling, and information management. These functions again are divided into strategic, tactical, and operative functions. The aim is to base the function model on an enterprise wide data model.

Function Oriented Integration Models

1. „Kölner Integrationsmodell (KIM)" by *Grochla* [Groc74]. In the mid 70s *Grochla* developed at the University of Cologne/Germany the „Kölner Integrationsmodell (KIM)" (Cologne Integration Model). The KIM is an integrated total model for planning, realization and control tasks of information systems. Data-proceeding jobs are seized within and between enterprise areas procurement, logistics, production, sales etc. The model consists of a literal and a graphical part.

   In the literal part a task description list, a channel list and a connector list are separated. The channel list describes how the tasks are combined and work together, the connector list brings tasks together with others. The graphical part exists of models for the named tasks planning, realization, and control. The KIM is very detailed and extensive. The focus is not on databases but on functions in the enterprise information system environment.

2. Integrated Information Processing by *Mertens* [Mert95]. *Mertens* divides in his model of integrated information processing (meanwhile the 11th edition) horizontal and vertical integration (Figure 3). The main focus for the horizontal integration aspect (which is more interesting in this context) is the value chain of an enterprise. Therefore he mainly deals with the coherence of processes (e.g. procurement) and does not start with an integrated enterprise-wide data model. His model is defined as text, in a function tree, a data flow plan, and in different tables.

   Although he points out that there is a risk for interrupting business processes when taking a view from departments or sections of an enterprise (and not from the process point of view) [Mert95, p. 4], *Mertens* deals with the functions within areas like R&D, sales, procurement, logistics, production etc. His model is likewise a total integration model and comprises the whole organization.

Figure 3: Total concept of the integrated information processing (Mert95)
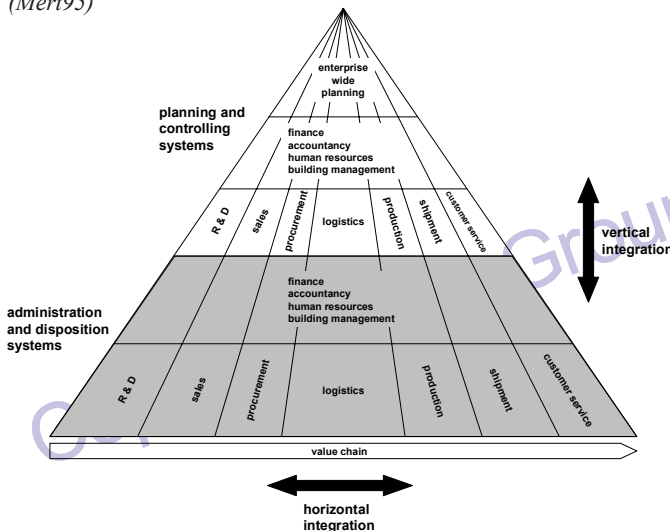
Other Approaches for System Integration

Integration is an implicit issue within the various approaches in business process (re-)engineering – especially when this topic is discussed from information technology's point of view. One of this approaches, *Scheer's* Architecture of Information Systems (ARIS) was discussed in a previous section. During the last decade a wide range of other business process engineering approaches were discussed. As most of them have a subset of the power in expressiveness or range of models, the following discussion will focus on SOM (Semantic Object Model), which is object-oriented and is different from the other approaches.

Introduced by Ferstl and Sinz in 1990 [FeSi90], the SOM is another approach to model business processes. It is targeted on enterprise planning, business process modeling, and the specification of business application systems. Enterprise planning is done to identify all relevant objects within the companies universe as well as outside. Then these objects are used as a base for the more detailed specification of the business process model from the behavioral perspective. The processes are split up in transactions between two objects. Transactions control and execute exchanged services and messages.

The business process model can be modeled from different perspectives: the interaction model specifies the flow between objects and the process event scheme shows the events related to transactions in a way similar to petri nets.

Similar to ARIS there is a special modeling tool available for the SOM approach. In contrast to ARIS it keeps only one meta model in its repository and all graphs are generated from that model. Also every modification made from one perspective will modify the central model and accordingly the other perspectives [FSAHM94]. Also based on this central model the third modeling goal in the approach will be supported: the specification of business application systems to ease it's implementation.

As most business process (re-)engineering approaches SOM is aimed to build a new application system from scratch. The tools accompanying these approaches are helpful to reduce the complexity of the according data, function, and process models resp. to handle them. But the resulting application is still complex and a closed one. When further applications have to be integrated, the right solution from this point of view is to develop a new application.

EAI SOLUTIONS PROVIDED BY THE MARKET

The heterogeneous environment of information systems leads to both many integration problems and big challenges for system architects. Historically built information infrastructures collide with claims made by current e-business needs. Related business processes are divided on different islands of hardware and software. EAI is demanded when business processes on different information platforms must be contracted. Currently, breaks in media and inappropriate human-to-machine interfaces are a standard in today's information environments. First steps of solving the problems lead to point-to-point connections between single applications. But the massive need for integrated systems in e-business can not be met with m*n single connection points. Experts say that after the database segment and the enterprise resource planning (ERP) segment EAI-applications will be next to become a big market of standard software tools [Nuss01].

EAI-tools link different data sources and applications at the application or process level. They portray business logic through process modeling and workflow control less than connecting systems technically or semantically on the functional or data level. They present a dedicated transaction platform either as a central hub-and-spoke model or as a distributed bus-oriented approach. [Knap01]

Integration can be done at three different layers of applications: on the presentation layer, on the function layer, or on the data layer. Integration at the *presentation* layer is quite simple because the respective information is put together at the user interface. The user is

able to work with different programs at one interface. It is possible to add features like workflow or validation rules. The disadvantage of this solution is the system performance because of the additional software layer upon existing programs. Implemented functions can be used further when the integration is made at the *function* layer. The applications communicate on interfaces and either exchange data or call a function. Integration at the *data* layer uses the data sources directly. This method is useful when data must be exchanged between different applications. Problems arise in understanding the data models and the business logic. Furthermore the interfaces must be adapted when changes at the data models are made; the administration tasks are quite intensive.

EAI and middleware are often used synonymously. But as an expert points out, EAI actually is the logical development of known middleware concepts. The middleware works on a lower level of abstraction either synchronous or asynchronous and is message oriented or transaction oriented. However, the method of EAI is more likely comparable to tools from the extraction-transformation-loading (ETL) area in data warehousing. Both provide the unite of data. [Knap01]

> „[...]middleware is a general term for any programming that serves to ,glue together' or mediate between two separate and usually already existing programs." [WiAu01] Many software producers offer their tools under the term middleware, whereas their marketing departments talk about EAI-solutions. However, the market can be subdivided in five different categories [Ober01]:

1. Remote Procedure Call (RPC) Technique. With RPC functions on a remote computer can synchronously be started by using a network. This oldest middleware technology comes from the 70s and is used for most distributed object technologies.
2. Database Access Middleware. This middleware enables the access on remote databases. Today, standards like Open Database Connectivity (ODBC) or Java Database Connectivity (JDBC) are accepted and supported by most databases. Database access middleware is suitable for access at the data layer only.
3. Message Oriented Middleware (MOM). MOM is used for the exchange of messages between applications. This service is an asynchronous method for integration. Each participant can be a sender or receiver, and the middleware plays the role of the post. Messages can be send and received either with the messaging method or through defined interface. Products are e.g. MQSeries by IBM, MSMQ by Microsoft, Java Message Service (JMS).
4. Distributed Object Technology (DOT). With the DOT objects are distributed in a network and can be called by defined interfaces. Today's applications often work with object oriented languages, so DOT plays an important role for EAI. Problems occur in the complexity of the interfaces. Products are e.g. DCOM by Microsoft, CORBA by the OMG, Java2 Enterprise Edition (J2EE) by Sun.
5. Transaction Oriented Middleware. In this technique transaction processing monitors (TPM) handle the communication of different applications. TPMs are the central node of the application logic and are named as the first generation of today's application servers. The procedures are based on single transactions. TPMs ensure the integrity of the transactions and manage the resources. Products are e.g. Tuxedo by BEA, MTS by Microsoft, CICS by IBM.

Different methods are often used in combination. The aim of using middleware is to generate consistent data structures in information systems. Moreover, EAI integrates business processes *and* data but not only data like middleware does. On the other hand middleware is basic for EAI-tools. While middleware does not provide a central control instance and leaves the business logic in the applications, EAI is process oriented and manages business logic and processes in a central authority. Uncoupling technology and applications enables the reuse of the single parts.

EAI ranks on a higher level than the single application or information system (and therefore middleware) does. The challenge of EAI is the aim to integrate existing systems on the one hand and on the other hand to define the business logic on a meta level (from technol-

ogy point of view). Starting point in EAI-projects must be the existing business logic and process connections realized in the applications. Hence, EAI primary is technology driven and business driven secondary.

COMPARING EXISTING AND NEW INTEGRATION CONCEPTS WITH CURRENT REQUIREMENTS

The previous chapter showed that vendors meet the technological challenges and that the scientific approaches concentrate on the conceptual view.

As stated before the scientific approaches are too simple to be helpful: building a new integrated application is a clear solution; but when facing limited time and budgets this integrated application is too complex to be a useful solution. Additionally, when integrating further applications the answer cannot be to build a new integrated one when considering limited time and budgets.

Therefore all the scientific solutions mentioned above are helpful to address some of the complexity problems. They help to handle the complexity of data, function, and process models by structuring them and – partly – offering the tools to do that in a structured and well documented way. But they are useful to document existing structures only. They are not connected to metadata repositories controlling ETL- or messaging tools. At least, all documentation work has to be done twice.

Table 1: Integration issues and their coverage in some research approaches

		Vetter	ARIS	Rau	KIM	Mertens	SOM
Integrated models for	Data	•	•	•			x
	Functions		•	•	•	•	•
	Processes		•			•	•
Metadata	Management		x			x	x
	Tools		x				x
	Methods		x			x	x
IT	Planning		x			x	x
	Maintenance		x			x	x
(•: fully covered, x: partly covered)							

The more theoretic approaches are helpful, but not satisfying as they do not provide methods to model different applications and the relationships concerning data, functions, and processes between them. They do not provide methods for an integrated interface management, integrated metadata collection and structuring, nor means to assure data quality, at least not in an integrated way The process models for these approaches fit to ordinary application development and often rely on prototyping, which is not adequate when the functionality for data exchange is not implemented at the required full extend.

The solutions offered by software vendors are focused on technical issues only. They offer tools to convert data between various formats, to exchange data between different kind of data storage systems, and to establish communication between older legacy and newer internet applications. Also, they offer tools to manage metadata and to interchange it with the operational systems as well as with other tools.

On the other hand they do not offer tools to model the integration. Data quality is recognized as an important problem, but the organizational side of this challenge is not addressed. The same is true for metadata management or the planning and maintenance of an IT

landscape of a company. Most tools cover one of the topics mentioned only, few more than one. And no tool or integrated set of tools covers everything mentioned.

Theoretic as well as vendors' solutions address only parts of these complex problems. An integrated approach is still missing. The basic elements of this approach are to be developed in the further work in our research group. These elements will be discussed in the following section.

OPEN RESEARCH ISSUES AND OUTLOOK

An integrated approach to address the problems mentioned above must cover—among others—the following topics:

- Metadata management as described before.
- Data quality management as described before.
- Data standards (XML etc.) are to be chosen to exchange the data between various operational applications.
- Transformation models are required to transform exchanged data between these systems.
- Project management has to address chains of integration projects as integration is an ongoing task and cannot be done with doing single and separated projects alone.
- The complex information technology infrastructure must be structured and planned to increase the flexibility of the overall IT-system of a company. Infrastructure development has to follow the evolution of the companies business and process modells and not vice versa.

To address these issues existing methods and tools especially in the data warehousing field have to be reviewed. Some methods and tools can be enhanced, additional tools and methods have to be defined and proofed in projects.

Scientific research combined with a close cooperation with practitioners daily facing these problems in their companies will be the base for the work in our research group in addressing these topics.

REFERENCES

[Bark90] Barker, Richard: *CASE*Method™ – Entity Relationship Modelling*, New York et al. 1990.

[EITO01] European Information Technology Observatory (EITO) – European Econmic Interest Grouping (EEIG) (Ed.): *European Information Technology Overvatory 2001*, Franfurt/Main 2001.

[EmGr90] Emmerich, Wolfgang; Gruhn, Volker: *Software Process Modelling with FUNSOFT Nets*, Software-Technology Memo No. 47, Dortmund 1990.

[FeSi90] Ferstl, Otto K.; Sinz, Elmar J.: *Objektmodellierung betrieblicher Informationssysteme im Semantischen Objektmodell (SOM)*, in: Wirtschaftsinformatik, 32 (6), 1990, pp. 566-581.

[FeSi94] Ferstl, Otto K.; Sinz, Elmar J.: *From Business Process Modeling to the Specification of Distributed Business Application Systems – An Object-Oriented Approach*. Bamberger Beiträge zur Wirtschaftsinformatik Nr. 20, Bamberg 1994.

[FSAHM94] Ferstl Otto K., Sinz Elmar J., Amberg Michael, Hagemann Udo, Malischewski Carsten: *Tool-Based Business Process Modeling Using the SOM Approach*. Bamberger Beiträge zur Wirtschaftsinformatik Nr. 19, Bamberg 1994.

[Groc74] Grochla, E. und Mitarbeiter: Integrierte Gesamtmodelle der Datenverar-beitung. Hanser, München 1974.

[Knap01] Knapp, Christoph: Intelligente Verbindungen - Mit EAI Prozesse und Workflow übergreifend steuern, in: Computerwoche Extra, No. 02, 23.03.2001, p. 11

[MeHo92] Mertens, P.; Holzner, J.: Eine Gegenüberstellung von Integrationsansätzen der Wirtschaftsinformatik, in: Wirtschaftsinformatik 34 (2), 1992, p. 5f.

[Mert66] Mertens, Peter: Die zwischenbetriebliche Integration der Datenverarbeitung im Einkaufs- und Lieferwesen, in: Zeitschrift für Datenverarbeitung, No. 4, 1966, p. 207.

[Mert95] Mertens, Peter: Integrierte Informationsverarbeitung 1, 10. ed., Gabler, Wiesbaden 1995.

[Nuss01] Nußdorfer, Richard: Der Markt hebt langsam ab - Enterprise Application Integration in Deutschland, in: Computerwoche Extra, No. 02, 23.03.2001, pp. 4-7

[Ober01] Oberdorfer, Roland: Allround-Adapter – EAI: Ordnung in Unternehmens-anwendungen, in: iX, Magazin für professionelle Informationstechnik No. 5, 2001, pp. 136-139.

[Rauh90] Rauh, Otto: Informationsmanagement im Industriebetrieb – Lehrbuch der Wirtschaftsinformatik auf Grundlage der integrierten Datenverarbeitung, Herne, Berlin 1990.

[Sche91] Scheer, August-Wilhelm: Architektur integrierter Informationssysteme. Springer, Berlin 1991.

[Sche95] Scheer, August-Wilhelm: Wirtschaftsinformatik Studienausgabe. Springer, Berlin 1995.

[Ubis93] UBIS GmbH (Ed.): *BONAPART TM – Model your business – Objektorientierte Unternehmensmodellierung mit BONAPART™*, Berlin 1993.

[Vett90] Vetter, Max: Strategie der Anwendungssoftware-Entwicklung – Planung, Prinzipien, Konzepte, 2. ed., Teubner Verlag, Stuttgart 1990.

[WiAu01] Without Author: „Content"; WhatIs.com, TechTarget.com, 2001

http://whatis.techtarget.com/definitionsSearchResults/1,289878,sid9,00.html?query=mi ddleware

[Wint00] Winter, Robert: Zur Positionierung und Weiterentwicklung des Data Warehousing in der betrieblichen Applikationsarchitektur, in: Jung, Reinhard; Winter, Robert (Hrsg.): Data Warehousing Strategie — Erfah-rungen — Methoden — Visionen, Springer, Berlin et al. 2000, p. 127-139.

0 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/proceeding-paper/enterprise-application-integration/31901

Related Content

Power System Fault Diagnosis and Prediction System Based on Graph Neural Network
Jiao Hao, Zongbao Zhangand Yihan Ping (2024). *International Journal of Information Technologies and Systems Approach (pp. 1-14).*
www.irma-international.org/article/power-system-fault-diagnosis-and-prediction-system-based-on-graph-neural-network/336475

Ontology Evolution: State of the Art and Future Directions
Rim Djedidiand Marie-Aude Aufaure (2010). *Ontology Theory, Management and Design: Advanced Tools and Models (pp. 179-207).*
www.irma-international.org/chapter/ontology-evolution-state-art-future/42890

The Effects of Sampling Methods on Machine Learning Models for Predicting Long-term Length of Stay: A Case Study of Rhode Island Hospitals
Son Nguyen, Alicia T. Lamere, Alan Olinskyand John Quinn (2019). *International Journal of Rough Sets and Data Analysis (pp. 32-48).*
www.irma-international.org/article/the-effects-of-sampling-methods-on-machine-learning-models-for-predicting-long-term-length-of-stay/251900

Location Based Context-Aware Services in a Digital Ecosystem with Location Privacy
Peter Eklund, Jeff Thom, Tim Wrayand Edward Dou (2013). *Cases on Emerging Information Technology Research and Applications (pp. 151-175).*
www.irma-international.org/chapter/location-based-context-aware-services/75859

Maturity for Sustainability in IT: Introducing the MITS
Martijn Smeitinkand Marco Spruit (2013). *International Journal of Information Technologies and Systems Approach (pp. 39-56).*
www.irma-international.org/article/maturity-sustainability-introducing-mits/75786