5-16-2003

Managing Information System Integration Technologies--A Study of Text Mined Industry White Papers

Balaji Ravindran
University of New Orleans

MANAGING INFORMATION SYSTEM INTEGRATION TECHNOLOGIES-A STUDY OF TEXT MINED INDUSTRY WHITE PAPERS

A Thesis

Submitted to the Graduate Faculty of the University of New Orleans in partial fulfillment of the degree of

Master of Science in Engineering Management

by
Balaji Ravindran
B.E., Bangalore University, 1999

May 2003

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my advisor Dr. Sathiadev Mahesh, for giving me an opportunity to work in this field and guiding me throughout the course of my thesis. I also wish to express my sincere thanks to the MIS professors at UNO whose help made this work fruitful. I would like to take this opportunity to thank Prof. Cherie Trumbach, Prof. Kurt Satter, Dr. Brett Landry, Dr. Greg Elofson, and Prof. Will Lannes for assisting me at various stages of this thesis. I also thank *ebizq portal* for providing invaluable resources on this topic.

I dedicate this thesis to those 'Enterprise Integration' projects that have failed during implementation stages over the past few years. I hope this work would be of some use for the future integration project team members. I wish this would be of some value addition for managers in the process of efficient decision making through industry white papers.

TABLE OF CONTENTS

ABSTRACT

Industry white papers are increasingly being used to explain the philosophy and operation of a product in marketplace or technology context. This explanation is used by senior managers for strategic planning in an organization. This research explores the effectiveness of white papers and strategies for managers to learn about technologies using white papers.

The research is conducted by collecting industry white papers in the area of *Information System Integration* and gleaned relevant information through text-mining tool, *Vantage Point*. The text mined information is analyzed to provide solutions for practical problems in systems integration market. The indirect findings of the research are *New System Integration Business Models*, *Methods for Calculating ROI of System Integration Project*, and *Managing Implementation Failures.*

LIST OF FIGURES

LIST OF TABLES

Chapter 1
Introduction

WHAT IS A WHITE PAPER?

In information technology, “a *white paper* is often a paper written by a lead product designer to explain the philosophy and operation of a product in a market place or technology context.” [Smith, D. Tom, 2002] It is a not too-detailed technical explanation of an architecture, framework, or product technology. Typically, a white paper explains the results, conclusions, or construction resulting from some organized committee or research collaboration or design and development effort. Whitepapers are the first step in the process of converting a technological dream into a reality.

MODERN TECHNOLOGY MEDIUM COMMUNICATION

There needs to be cooperation in an organizational context grounds on communication. Malone’s approach describes the relationships between coordination, cooperation, communication and communication media in a top-down manner. There has to be an appropriate communication medium which would improve communication and thereby indirectly contribute to cooperation for the introduction and adoption of new technologies for system integration.

"Asynchronous and synchronous operations are complementary subparts of large tasks or activities. A meeting proceeds in a largely unstructured way, but it can contain islands of structured synchronous operations. This calls for integrating support for structured/unstructured activity on the one hand and for synchronous/asynchronous activity on the other." [Schoder, 1996]

USE OF WHITE PAPERS BY MANAGERS

Klare (2000) says a writer should ideally have a detailed picture of his audience. An author should have information about reading, educational and intellectual levels, background or previous experience with author's topic, interests, level of motivation, and voluntary or involuntary attention to the material. The audience has to be figured out using readability formulae. Doing this before the writing is started will make the writing more mechanical and take away the individuality of style. The concern should be on organization and emphasis. If the writer decides he may not reach his/her audience, revision should be called for. The emphasis would be on the structure of writing; the content of the material is decided by the author, and it should remain his/her concern.

Hargis (2000) ranks the levels of readability in descending order as legible, interesting, doable, comprehensible, translatable, and learnable. Whitepaper readers are normally part of a specialized audience. Klare (2000) acknowledges that if the writer is interested only in a small, specialized, highly educated audience, the principles of readability presented may not be of great concern. But the argument is that the users have to deal with the complexity of the topic. New users seemed to be much less tolerant of material that is hard to understand and use. Research on cross-cultural characteristics such as individualism and collectivism (Hofstede, 1980), high context and low context communication (Hall, 1973), and universalism and particularism (Trompenaars, 1998) is now being applied to technical communication (Honold, 1999; Thatcher, 1999).

"IBM developed an early grammar and style checker called EPISTLE, which was intended to help stamp out gobbledygook as well as poor constructions in business communications, later used by technical writers." [IBM Corp., Readability and computer documentation, 2000]. The two major approaches to the content are task orientation and minimalism. The principles of task orientation involve understanding the audience, specifying the user's tasks, and organizing the information around sets of instructions with

supporting conceptual and reference information. Researchers at the IBM Watson Center found out that the documentation needed to be reengineered; cutting away much information and adding information to help users get started quickly on realistic tasks and to give them pointers when they made mistakes. This was called as “minimalism”.

The six characteristics of quality improvement model are accuracy, appearance, extraneous information, missing information, readability and usability (IBM, 1985a). Information on the web has three parts that make its three-dimensional nature: architecture, interface, and interaction. “Architecture is the structure and organization of information. Interface is the element between the user and the computer that conveys the architecture. Interaction is action the users assign to the elements based on how they perceive.” [Zibell, Kristin, 2000]

AUDIENCE, PURPOSE, AND CONTEXT OF WHITE PAPERS

“The user is whoever we might imagine using our texts. Any nontrivial understanding of audience has to go beyond simple demographics to what the audience is doing and what goals they expect to meet.” [Spinuzzi (1999)] For instance, an author writes a manual describing how to use Microsoft Word. He assumes that the end user (the imagined audience) and the author share the goal of teaching the user how to (among other things) create new documents, style, use styles, and avoid calling technical support. “But the user may have aims that the author does not share, such as writing Word macro viruses or calling technical support people at their homes at 3 A.M. The author of the whitepaper does not document this useful information because although it meets the user’s goals, it contravenes the goals of his/her organization.” The context of the whitepaper relies on the user’s purpose. Business managers use the whitepapers to make future decisions and write strategic plans, while analysts might find the same whitepaper as an update of technological ideas, and would consider the same information for further technical improvements.

GLEANING INFORMATION FROM WHITE PAPER

Systems that allow groups to collaborate over computer networks provide support to coherent interactions and ideas that enable groups with a shared aim to make progress toward a common goal. Text mining assists in gleaning over huge amount of unstructured texts to find patterns, relationships and commonalities between keywords.

Chapter 2
Information System Integration

A CONTEXTUAL ANALYSIS

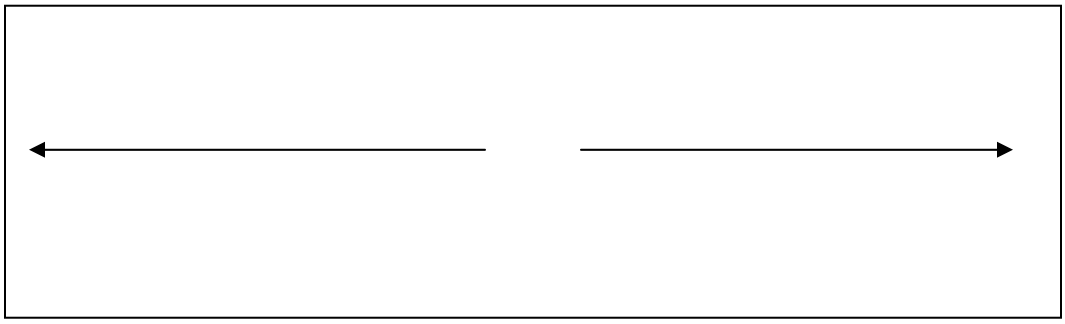
The aim of information system integration is to increase the profit of the enterprise by integrating disparate information systems for automatic exchange of business information. For each enterprise to get specialized on a specific function there is a need for inter-enterprise integration, as well as intra-enterprise integration. [Kalakota and Whinston, 1996] The important factors to achieve inter-enterprise integration are standardization and digitalization.

To have an idea of how an enterprise functions from different points-of-view would be of great advantage to achieve good system integration. Leem, Shin introduce reference architecture models such as CIMOSA (Computer-integrated manufacturing open system architecture) framework which follows a top-down and process-oriented structural approach. It defines four different views concerned within the enterprise- function, information, resource, and organization. Out of which there are three major sections. They are functional, behavioral and structural.

Integration Patterns

Different kinds of integration problems need different approaches for a solution. For instance, integration using synchronous communication methods demands a different approach than that of asynchronous communication. William Rue defines the main integration patterns as monolithic, data-consistency, multi-step, composite and autonomous distributed.

Using traditional programming methods, self-contained systems are built which are known as “monolithic” applications. This pattern binds the applications modules tightly to each other, which in turn makes the modifications more expensive and complex. Eg: Legacy systems.



Monolithic Data Multi-step Composite Autonomous
n Consistency Process Applications Distributed

Figure 2.a: INTEGRATION PATTERN

Data consistency pattern follows the coordination of shared data by exchanging data between application and systems. Multi-step integration patterns are illustrated by applications using asynchronous exchange of transactions such as the procurement application integrating requisition, approval and purchase generation as a sequence of transactions. Composite application integration patterns are used by those systems that implement their behavior through method calls to external application services. Eg: a cargo application verifying credit, calculating freight costs, and estimating delivery date by invoking external credit rating, tariff calculation and delivery routing services. This kind of integration is popular in applications where synchronous communication and near-end time response are required. Autonomous distributed pattern is characterized by the application the combination of application services. It enables integration of services to occur at runtime. UDDI, EADL, SOAP (Web services framework) provide a basis for such a dynamic integration approach.

[Puschmann, Alt et al, 2001] There are three kinds of integration approaches- homogenous with one instance, homogenous with several instances, and heterogeneous. Heterogeneous integration has been classified into four layers. Process level integration

deals with the pragmatics of two systems. Object integration deals with semantics, data level deals with syntax, and OSI-7 layer model integration deals with communication services.

INFORMATION SYSTEM INTEGRATION-FRAMEWORK

Information System Integration is the linkage of systems that are developed to streamline the business processes so that they work and function as a single entity. The diverse systems could be spreadsheets, structured and unstructured databases, web pages, e-mail and presentations. The challenge faced is to target the customer requirements in real-time.

System Integration is accomplished by defining which internal data items and processes should be shared, or made available, to other services. Making the data and processes of a particular service available to external influences is the first step in allowing it to integrate with other business functions and services. Data and processes are exposed to the integration layer by using interfaces or adapters. It is these adapters that isolate the service or application, and therefore allow supplier or user systems to continue to develop and change independently. The isolation is required to avoid impacting the integration solution and the community as a whole.

Due to a competitive global market there are a large number of new entrants with radically different business models that can effectively use disruptive technologies, a net savvy marketplace that seeks new economy. Business needs to improve the velocity of intra and inter business process. It must offer accurate information at the right time, allow for inter-unit collaboration (sub-systems within unit and external elements), make system/platform transparent to user, adapt to new business models quickly (low cost, effective, fast). [Mayer & Painter].

Integration within the enterprise happens when applications are shared between different functional areas within an organization to optimize business process flow. [Synergy Connex] Intra enterprise integration is achieved through a common database model and centrally designed systems, or proprietary

connectors and buses linking different systems. This integration may be at the data level or at the process level.

Standardization and digitalization helps inter enterprise integration by supporting the integration of independently developed information systems for data or process integration. [Leem, Seong et al. 2002] There is a tightened data flow between an organization and its trading partners, vendors, or suppliers within a value-chain.

A completely new system can be developed to achieve seamless integration between many business processes. Such a system would be based on a common database, using an entity wide data dictionary, and processes using a common model. The dynamic nature of business, especially in the net era, means that any business model needs to adapt change continuously, or become rapidly obsolete. New business processes are necessary and established ones need to change. The need for systems integration is not eliminated by the selection of an industry wide IS solution.

Data integration approaches transfer data between systems, and maintain the integrity and security during this process. Process integration links processes in one information system to processes on another information system. In addition to integrity and security of the process, interchange needs to assure load management of processes in the integrated system.

Data Integration

Data in one system needs to be transferred to other systems to complete business processes. Asynchronous data transfer mode (ADTM) is a high speed, connection oriented switching and multiplexing technology that uses short, fixed-length packets (called cells) to transmit different types of traffic simultaneously, including voice, video, and data. [DePrycker 1993, Newman 1994, Fischer et al. 1994, Haendel et al.1994, Vetter 1995, Kim and Wang 1995]. Asynchronous data transfer is a connection-oriented technology where a sender first establishes a connection with the receiver. However, unlike circuit switching, in which a physical circuit is established between the sender and the receiver and reserved for them for the entire duration of their communication

session, in ADTM technology, a virtual circuit is established between the sender and the receiver. That is, ADTM does not reserve the path for one user exclusively. Any time a given user is not occupying a channel, another user is free to use it. Data is downloaded and stored on a storage device that is typically accessed via the network. This makes the stored data accessible by both systems. Access control lists need to be set up to limit access to the transferred data to authorized users or applications. Storage area networks speed up this transfer by streamlining the access control process. The format of the downloaded data needs to meet the requirements of the application to which it is uploaded. This format could be a lowest common denominator format such as a fixed width text file or a comma delimited text file. In order for the transfer to be effected the uploading application must have a published uploading format that specifies the order and type of data and the downloading application must have the required data in the specified format. When the downloading application is either incapable of delivering the data in the specified format or it is difficult to modify the application to deliver the data in the specified format, a data transfer script can be developed to make the necessary format changes. These format changes could be simple data order, element demarcation, data type changes, data value changes, or data storage format modifications. When the data type or storage allowed for the uploading application is unavailable on the source data set, the transfer script need to make intelligent translations of source data. Similarly data value changes need programs that can make the necessary modifications. Since asynchronous data transfer is typically a batch process, many transactions are transferred in each cycle, and the uploading application must be capable of executing a batch update.

The development of asynchronous data transfer scripts between applications begins with a definition of the format required for the destination application, a study of the data available in the source application, and the development of a suitable transfer script. The batch data transfer process becomes a new application that links data in the two applications asynchronously, and is periodically

executed. In addition, separate processes to download and store, and to retrieve and upload are executed on the source and destination applications. The three processes need to be sequentially executed. The processes often exist on different platforms. The source and destination applications typically run of completely separate systems, and the data transfer application may run on either platform, or even on a totally independent platform. Integration of these three processes may be achieved by process notifications that trigger each step, a synchronized system where each process is run at a specific time, or by establishing a communication capability between the data transfer applications.

These data transfer applications need to be maintained when a component of either application is modified through a maintenance process, and results in a change in either delivered data from the source, or input format for the destination. Changes in the execution of processes at either end could also force modifications to the transfer system.

The data transfer process needs continuous monitoring for quality through built in audit checks and maintenance when notified of changes to either system or after a notification of an error. Each one way asynchronous transfer of data requires execution of transfers, monitoring of performance, and updates to the transfer script when changes occur in either application. Hence transfer of day from one application to many (n) applications requires management of n data transfer systems, a two way data transfer requires 2n data transfer systems. Transfer of data between n systems requires 2(n)*(n-1)/2 data transfer systems to be manager. This rapid escalation of data transfer complexity overwhelms IT departments. A standard for data transfer reduces the number of data transfer systems to be developed and maintained. For example the adoption of an industry wide standard for input of transaction data and the ability of applications to output data to this standard reduce both the number of translations and the likelihood of error in these transfers.

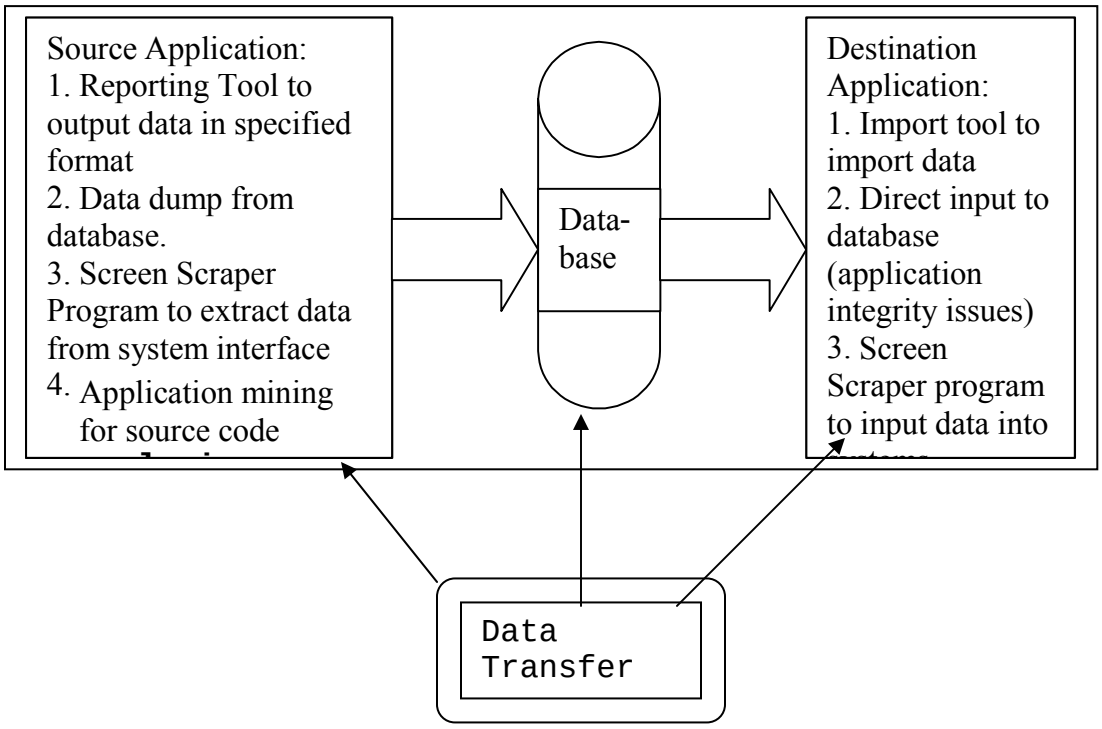


Figure 2.b: DATA INTEGRATION

In file-level data transfer model, when an operation requires file data to be transferred across the network in either direction between a client and a server, the whole file is moved. [Satyanarayanan et al. 1985]. Transmitting an entire file in response to a single request is more efficient than transmitting it page by page in response to several requests because the network protocol overhead is required only once. There will be reduced server load and network traffic due to better scalability. Disk access routines on the servers can be better optimized if it is known that requests are always for always for entire files rather than for random disk blocks.

In Block level transfer model, file data transfer across the network between a client and a

server take place in units of file blocks. A file block is a contiguous portion of a file and is usually fixed in length. For file systems in which block size is equal to virtual memory page size, this model is also called a page-level transfer model. It does not require client nodes to have large storage space, and also eliminates the need to copy an entire file when only a small portion of the file data is needed. Therefore, this model can be sued in systems having diskless workstations. It provides large virtual memory for client nodes that do have their own secondary storage devices. E.g.: Sun Microsystems NFS [Sandberg, 1987].

In Byte-level transfer model, file data transfers across the network between a client and a server take place in units of bytes. This model provides maximum flexibility because it allows storage and retrieval of an arbitrary sequential sub- range of a file, specified by an offset within a file, and a length.

The record-level data transfer model is suitable for use with those file-models in which file contents are structured in the form of records. The Research Storage system (RSS) [Gray 1978, Gray et al. 1981], which supports complex access methods to structured and indexed files, uses the record-level transfer model.

EDI standards provide strictly defined industry wide formats for fixed width text files. XML tags allow source programs to output tag identified content that can be parsed by common programs for input into destination applications. This permits data transfer programs to be generic and transfer data from a wide range of input sources, making the management of data transfer programs more convenient. Industry wide adoption of a XML tagging scheme will maximize the benefits from XML. XML tagging allows data transfer for more complex data than the alphanumeric data usually handled by DTS. For example e-mail files with attachments, presentations, and rich media files can be transferred more conveniently across systems if they are effectively XML tagged.

Synchronous data transfer systems transfer data directly from one system to another. The major challenge of synchronous data transfer is the need to manage a large number of transfers of data, typically

one record or less at a time. The volume of data transfers makes manual or semi-automatic transfers impossible and automated data transfer is essential.

Interface Integration

A screen scraper is a tool used to extract data from web sites, and consists of two main parts. The first is a proxy server that can be run locally in order to view in plain text the contents of both HTTP and HTTPS requests and responses as they pass between a web browser and server. The second is an engine that can be easily configured to extract information from web sites, handling such tasks as authentication to a site, following redirects, and automatic handling of cookies. Special patterns, which can optionally make use of regular expressions, are used to identify and extract data. An embedded scripting engine is included so that extracted data can be manipulated, written out to a file, inserted into a database, etc. screen-scraper is written in Java, and has been tested on Mac OS X and various flavors of Linux and Windows.

Application mining provides detailed technical information about the real internal structure of the applications. Application mining is based on the combined use of source code analysis technology, the storage of resulting metadata about the applications in a repository, and on visualization and reporting systems that then use this metadata, for providing the information on the internal structure of existing applications.

There could also be combination of system integration such as interfaces to do processes.

Process Integration

Standards-driven packaged wrappers first appeared in the early 1990s for the integrated enterprise, with client-server architecture becoming successful. This became the foundation of process re-engineering and integration, with a business change. [Watson, Schneider et al] Process integration coordinates responses to business events, determines sequence and execution conditions, tracks process status and exceptions, and generates alerts. Process

integration solutions have become commonplace in current markets, but current vendor emphasis is on proprietary offerings, not standards. In the past, approximately 80% of business processes could be standardized. [Watson, Schneider et al] It is critically important to the automation of both internal and inter-enterprise systems, so both vendors and users will come under increasing pressure to standardize process models and protocols. Business processes are extending beyond the organization. The focus is still on the customer, but instead of immediate customers the business looks at the end consumer of the supply chain.

Earlier process integration lacked support for implementing interfaces to automated systems without expensive, time-consuming programming. In recent days, new solutions are emerging which support both business process modeling and interface-focused information flow modeling, without low-level programming. By merging these two approaches to the multi-pattern, organizations gain the ability to capture and model process integration requirements at the business level, including long-duration processes and processes that span organizational and system boundaries, like collaborative product design. But they also realize faster implementation of interfaces to back office applications, web applications, enterprise data stores, partner systems, and IT resources, and the ability to monitor and manage the operational result using business-level process models. [Mercator et al, 2002]

Architecture is a key consideration in data transformation. Although point-to-point interconnections can do the job, much of the transformation logic gets duplicated in each link, and the maze of links quickly becomes unmanageable. The alternative is the hub-and-spoke architecture. Within the hub, there is the question of whether the data is transformed on the fly or with intermediate staging to files. [Strehlo et al] On-the-fly transformations won’t scale; on the other hand intermediate staging in a data store imposes significant performance overhead. To do so, architecture needs to allow segmentation of the process and control over the location where extraction, transformation, and load (ETL) processes can run. Batching of processes to single server have

to be done with the possibility of intermediate data storage, or partition them to run in parallel on a set of servers or serially in a pipeline of servers, as operational and load-balancing needs dictate. The successful completions of total process have to be ensured.

Application Integration

Application integration is the corner stone of modern IT strategy, and therefore of modern enterprise strategy, particularly for organizations that operate in dynamic, competitive markets. Hasselbring defines three dimensions of Systems Integration problem. They are autonomy, heterogeneity and distribution. [Hasselbring et al 2000] Autonomy of component systems may exist in design, communication and/or execution. Systems integration tries to enforce standards to reduce autonomy, while it is necessary to allow developers meet sub-unit requirements. Organizational changes are required to reduce autonomy, such as implementing the use of COM+ or EJB across one or multiple organizations. Distribution is the range of calling conventions, often occurring due to legacy systems. Proxy services (also used in firewalls) present the appearance of an integrated system. The Remote Procedure Call (RPC) approach eg: IIOP (CORBA) extends RPCs to an original setting.

Heterogeneity arises from different programming and data models. For example address as fields in one case, object in another, wrappers may be used – often not known until basic access problems are solved. Enterprise Application Integration (connectors) addresses semantics of the link while Middleware integration addresses syntax on the link (Hasselbring). Off late OMG has addressed some business logic (semantic) links as well moving middleware links to high level services. Hasselbring’s diagram reflects three layers – Business Architecture, Application Architecture, and Technical Architecture. Integration at the Business Architecture level is based on inter-organizational process. Similarly at Application Architecture level it is EAI, and at Technical Architecture level it is Middleware integration.

As integration needs increase, a clear framework is necessary for developers to effectively and efficiently build robust applications that span organizational boundaries. Typically, in such an environment, developers create applications not by writing large blocks of code, but by selecting the necessary components for the application, and extending the services available from these components. The term application framework is used to describe the available components (directory of services), the capabilities and responsibilities of these components (descriptions), and the interactions between these components (processes). For example the IBM Application Framework for e-business described by Flurry and Vicknair, provides a comprehensive set of standards that allow for intra-organizational systems integration. Standardization in intra-enterprise environment is based on common data models, processes and rules (ERP, EAI).

SUMMARY

Data integration approaches transfer data between systems, and maintain the integrity and security of the data during this process. Process integration links processes in one information system to processes on another information system. In addition to integrity and security of the interchange process interchange needs to assure load management of processes in the integrated system.

System Integration is accomplished by defining which internal data items and processes should be shared, or made available, to other services. Making the data and processes of a particular service available to external influences is the first step in allowing it to integrate with other business functions and services. Data and processes are exposed to the integration layer by using interfaces or adapters. It is these adapters that isolate the service or application, and therefore allow supplier or user systems to continue to develop and change independently. This isolation is required to avoid impacting the integration solution and the community of systems as a whole.

SCOPE OF THE RESEARCH

A wide range of systems integration technologies have evolved over the past few years. The challenge for business managers is to select and implement the appropriate technology based on the words of technologists. This means technology vendors need to communicate and market their services through proper channels. The main method of industry communication is through whitepapers published by technology companies. Also journal publications and news reports are used by managers to update their knowledge of technology. This research study focuses on the means of whitepaper communication of technology in system integration market.

THESIS RESEARCH MODEL

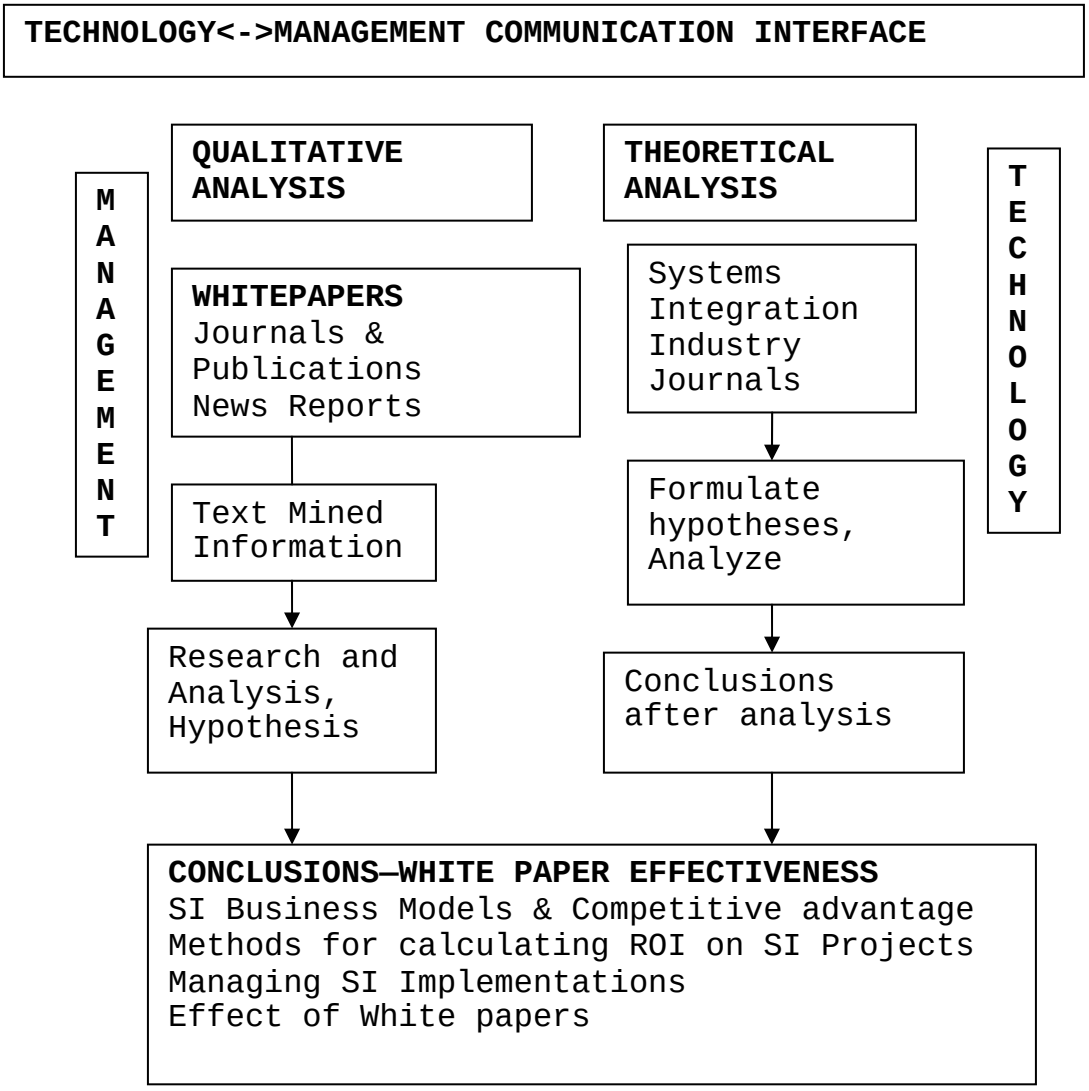


Figure 2.c: THESIS RESEARCH MODEL

RESEARCH QUESTIONS

1) How whitepaper communication is effective in system integration (SI) markets?
2) What are the new business models and the competitive advantages due to system integration (SI) technologies?
3) What are some of the methods of calculating the return on investments (ROI) of the system integration (SI) projects?

4) How to manage difficulties during Enterprise System Integration (SI) implementations?

Chapter 3
Literature Review

INTRODUCTION TO LITERATURE REVIEW

Informal and formal sources were accessed from the Internet and from the academic libraries. The information needs were expressed at the beginning before confining the area of research. Some of the needs were contact with industry professionals to provide answers for doubts concerning the systems integration technology, relevant research publications in systems integration and technology communication. To identify the component parts of information needs, a research model was framed by using mind map. This model was used to plot ideas for seeking further information. (Attached Chapter 1.5) Identifying the synonyms and keywords by the concepts already developed, a topic was chosen. The concern at this stage was about the communication between technologists and managers. As an engineering management student, the task was to identify the ways to make the communication effective. Whitepapers were identified to be the most popular modern mode of communication between the technologists and managers. Hence the area of research was confined to this level at this stage. The specific topic of this thesis is “Whitepaper communication in Systems Integration Markets”. The purpose of this research is to identify the effectiveness of industry whitepapers to the business managers for making decisions. A wide context, “Information systems integration” was chosen for research study as a business manager. No specific research on the same topic was already done. Hence, relevant materials of systems integration such as industry whitepapers, publications, news journals and articles were collected at the beginning of research, in the form of a digital library. Some individual authors were contacted for access to published and unpublished documents in the areas of system integration.

INFORMATION SEEKING AND RESEARCH METHODOLOGY

The whitepapers and industry reviews are gathered from the corporate websites and systems integration portals. The knowledge about the trends and technologies are accumulated prior to the analysis. The whitepapers and industry reviews are text mined for more relevant information by finding out the relationship between keywords statistically using the text-mining tool, Vantage Point. Vantage Point is a new text mining tool developed by Search Tech in collaboration with Georgia Tech University. The qualitative research is done with the gleaned information obtained from the mined texts. Finally, the conclusions are written based on hypotheses and supporting evidence. Hypotheses are written assuming the candidate in the position of business manager who has little knowledge about the field of 'systems integration'. Along the process of reading the whitepapers the subject is learned by the candidate as well as the effectiveness of the whitepapers is judged qualitatively, as against the quantitative results provided by the industry professionals on the same topic.

EVALUATION OF AVAILABLE INFORMATION

Technology Communication White Papers

Readability is one of the prime issues hinged on the success of communicating technologies through whitepaper. Klare, R. George in "The Measurement of Readability: Useful Information for Communicators" discusses the applications of readability principles and formulas. This research represents the interpretation of a survey data. Hargis, Gretchen in "Readability and Computer documentation" identifies that reader appropriateness of technical text involves readability. But he contradicts Klare's views by hypothesizing that there is no formula to measure readability.

In "Beyond Automation: A Framework for Supporting Cooperation", the researchers have used

Malone-Crowston layer framework for cooperation. This framework describes the communication in a top-down manner between cooperation, communication and media. A method is proposed to improve communication and thereby, indirectly contribute to cooperation.

Information System Integration Research Papers

Schneider and Watson published a report in the manufacturing journal at Louisiana State University on “Enterprise Systems and Business Integration”. This paper is mainly a follow-up research by Thomas Davenport (1998), “Enterprise Systems”, Harvard Business Review, and his book, “Mission Critical”. This paper finds the structure of new business models, popularity of collaborative business, and explains the types of emerging integrated business environment. The “information integration” part of this research is used as a predecessor of the research on “information systems integration”. Hasselbring defines the dimensions of Information Systems Integration as autonomy, distribution, and heterogeneity, in his publication “Information Systems Integration” in the “Communications of the ACM”, 2000.

Themistocleous, Irani (2001) analyzes the ERP problems through an Internet based survey. The conclusion of this paper is EAI could be the solution to ERP’s problems of integration. Whereas Puschmann, Alt evaluated the achievement of standardized integration architecture for different EAI systems. This is made possible through different integration approaches. The inter-organizational concept base results in enterprise integration by acting as essential ingredients for inter-organizational workflow modeling. Moor, Jeufeld in “Concept Integration Precedes Enterprise Integration” proposes ways to structure inter-organizational concept base. The methods to co-develop such concept bases from various enterprises are also illustrated.

THE STRUCTURE OF THE STUDY

The research thesis is structured in two steps. They are theoretical and qualitative analyses.

Theoretical Study

Text mining technology is used for gleaning relevant information from the large chunk of whitepapers on systems integration. This technology would provide some interesting information which could not be done manually, such as the factor maps, correlation maps and factor matrix. The relationships between the keywords would be of invaluable information. This relevant information is used for conducting a theoretical analysis based on previously published papers in the same area. The best practices of 'Information Systems Integration' would be interpreted from the whitepapers. The results would be found out for the specific thesis questions. The conclusion of theoretical analysis would reveal how effective the whitepapers are for the business managers.

Qualitative Study

Qualitative analysis is conducted based on the results of theoretical analysis, as done in Chapter-4. These results would provide the answers for research questions in the form of conclusions of the research done so far, which will be used for valuable judgments. Also the findings of qualitative analysis would be compared with that of quantitative analysis for concluding the thesis.

JUSTIFICATION OF THE STUDY

This research finding involves two ways (theoretical and qualitative) to verify the results before a firm conclusion. Hence the discrepancy in the results between these methods would be taken care of at the end of research. The common aspects of the results would be valued much higher and would be utilized for writing conclusions. Wherever there are significant discrepancies necessary steps would be taken to modify, or curtail certain points for perfection.

Chapter 4
Methodology

TECHNOLOGY AND STANDARDS IN INFORMATION SYSTEM INTEGRATION-A CONTEXTUAL DESCRIPTION

The standards of Information System Integration include information, format, and business processes. For a better information sharing and good business interface latest system integration technologies are essential. Behaviorally, an open-minded culture to share information is necessary which will increase the overall performance of the enterprise.

Some of the main technologies of Information Systems Integration are as follows:

- Enterprise Resource planning (ERP)
- Enterprise Application Integration (EAI)
- Enterprise Portals
- Integration with legacy systems
- Customer Relationship Management (CRM) Integration Communications middleware
- Java messaging service
- Mobile or ubiquitous integration
- Real-time infrastructure integration
- Web services
- E-business integration
- Enterprise Data Interchange (EDI)
- Knowledge integration,
- Supply-chain integration and
- Exchanges

Table 4.1 – FORCES AFFECTING SI TECHNOLOGIES

S. No.	Name of the Technology	Organizational environment	External forces affecting the progress
1.	Enterprise Resource Planning	Any organization	Organizational resistance to change
2.	Enterprise Application Integration	Any technological company	Top management support
3.	Enterprise Portals	Tech companies	TMS
4.	Legacy Integration	Org following traditional systems	Latest technologies and web
5.	Customer Relationship Management	Most of customer focused companies	Greater customer service and friendly management
6.	Communications Middleware	Hi-tech companies	Competitor’s product
7.	Java messaging services	Hi-tech organizations	Products from Microsoft and IBM
8.	Mobile Integration	Mainly communication organizations	Latest ubiquitous technologies, telecommunication policy
9.	Real-time infrastructure integration	All organizations	High cost and governmental interference
10.	Web services	All organizations	Lack of resistance by traditional tech companies
11.	e-business integration	All organizations	Networking technologies
12.	Enterprise Data Interchange	All organizations	New technologies-EDI is outdated
13.	Knowledge Integration	Tech companies	People rejecting expert systems
14.	Supply chain integration	Manufacturing	Poor supplier infrastructure
15.	Exchanges	Trading companies	Companies using phone and other modes heavily

TEXT MINING RELATIONSHIPS FROM FACTOR AND CORRELATION MAPS

The maps attached in the *appendix-A* are based on "text mining" of those records to study patterns of 'Information System integration'. Using Vantage Point software, factor maps and correlation maps were generated based on 81500 keywords. The main headings in the 'factor map' mostly come as the most central keyword of given clusters. The number in the parenthesis after each topic represents the number of articles within that factor. For each topic, sub-topics are the high loading keywords for that cluster. The node size reflects the relative frequency in the dataset of 250 white paper records. The placing of keyword is based on a Vantage Point proprietary 'Multi-dimensional Scaling' (MDS) routine. The connecting lines indicate relative degree of association as per its boldness, using a Path Erasing algorithm. The absence of a link suggests less or no association. In a correlation map each node represents one technology (as in this case). Proximity of technology and links between technologies show high correlation among inter-related usage of technologies.

From these maps, the user can interactively view any field for each node. For example, as the user moves the mouse pointer over a node, a pop-up box shows the list of sub-keywords.

ANALYSIS USING TEXT MINING TECHNOLOGY – BIBLIOMETRIC ANALYSIS

Vantage Point, a text-mining tool developed by Georgia-Tech was used to analyze what is being written about the information systems integration technologies. When the lists of abstracts and 250 whitepapers were fed into Vantage Point, it gave an interesting result of around 81500 keywords.

The factor matrix and co-occurrence matrix were listed by choosing particular information systems integration technology keywords. The factor matrix is a table displaying the factor loadings of all variables on each factor. Some of the important keywords were rules, CORBA, C++, HTTP, XML, exchange,

SAP, Java, MQ Series, MSMQ, impact, business models .etc., The mainly listed keywords from the factor map were EDI, strategy, automation, legacy applications, analysis, Interaction, any-to-any communication, transaction, reliability, CORBA, and dynamic. (More information on Appendix).

IMPORTANT KEYWORDS

- CORBA
- Rules
- C++
- HTTP
- XML
- Exchange
- SAP
- Java
- MQ Series
- MSMQ
- Impact
- Business Models

This factor mapping was done after a cluster analysis. The cluster analysis is a multi-variate technique with the objective of grouping responses or cases with similar profiles on a defined set of characteristics. There was a very high relation between the keywords automation, reliability and transaction. The set of phrases associated with these keywords were automation, enterprises, investments, impact, billing, supply chain, technology, market, success, process.

The keyword, “transaction” was more related with EDI technologies and legacy applications. CORBA technology was in isolation with any of the supply-chain related keywords. The keyword, “strategy” was linked to the main keyword of the factor map, “reliability” which is again inter-linked with “transaction” and “automation”. This shows the importance of strategy in the system integration.

The isolated but existing keywords were “interaction” and “any-to-any communication”. This means when talked about the traditional technologies like EDI or legacy systems, the modern day technologies exists in isolation. This is because of on-going integration and up gradation of age-old

technologies with new systems. This is the main role of information system integration.

From the auto-correlation maps, it is clear that the keywords, “ERP”, “suppliers”, “CRM” exist but in isolation. Another set of inter-related but isolated from the mainstream keywords are “marketplace”, “strategy”, “planning”, “communication”. This shows that there is a heavy link between strategic planning of organizations’ involvement with those of communication and information system integration technologies. Whenever these technologies are talked about in context with the market places, strategic planning comes into picture. Similarly, a very high link between “customers”, “investment”, “processes” and “markets” also occur. This depicts the relationship of quality and processes with customer relationship and supply-chain integration markets.

The auto-correlation map of CORBA shows a picture that is related technologically. Java is seen at the heart of the information system integration technologies. There is a heavy link between Java and server technologies. Some of the examples in this context would be web servers such as “Sun”. CORBA is highly related with “Application severs” and “Java technologies”. Java is related with web servers and CORBA with app servers. The keyword, “clients” shows a strong relationship to the word, “server” which means client-server applications are related to Java and CORBA technologies. The keywords, “databases” and “architecture” are related with Java and server technologies. This shows a weaker link. This may be a prediction that these technologies provide a robust architecture. Although the keyword, “operating systems” exists, it is seen to be in isolation. The most interesting part of this map was seen at the bottom of the map with two giants “IBM” and “Microsoft” related to each other and not with any of these technologies. The interpretation of which is IBM and Microsoft are the major competitors with Sun Microsystems in this field. IBM’s MQ-Series and Microsoft’s MSMQ are the major middleware products that are in competition with the middleware, CORBA.

Another correlation-map shows a greater relationship between legacy integration with application integration, which means that the old

legacy systems are upgraded to the latest applications through information system integration.

A MODEL OF THE INFORMATION SYSTEM INTEGRATION PROCESS

Standards for Business Processes already exist. Eg: BPML (Business Process Modeling Language) “BPML allows business processes to be managed outside the applications, which potentially will foster increased collaboration and innovation between enterprises” (Smith, 2001) BPML helps deal with higher level objectives by moving the business forward by better decision making. Emerging B2B protocols such as ebXML, BizTalk, RosettaNet are bridged together with IT infrastructures using BPML. A similar model or standard is necessary for ‘Systems Integration’ processes for the organizations to gain more values in corporate and technological decisions.

Chapter 5
Results

THEORETICAL ANALYSIS AND INTERPRETATION OF TEXT MINED INFORMATION

The relationships between keywords from factor maps are interpreted for further analysis. The thickness of the line between the keywords expresses the strength of relationship. Such relationship analyses are done for every set of keywords and selected white papers are read for interpretations.

Interpretation from White papers

"Strategy" and "Information System Integration Technologies"

Technological innovations have helped businesses prosper over the past few decades than ever. New possibilities arose by technologies have helped business organizations create a new paradigm shift. The keyword 'Strategy' is mainly related to the word, 'Outsourcing'. This could be explained by strategic outsourcing. "Strategic outsourcing is a process that enables company procurement managers to select the most effective suppliers to use for direct and indirect procurement while optimally allocating business among supply base"[Dalrymple, Kelly 2002]. The decision is based on a set of criteria such as best price, delivery time, location, volume discounts, deadlines, quality, quota restraints, and other factors. E-sourcing is the automation of that complex process using sourcing and collaboration software rather than traditional means of phone, fax, e-mail and letter.

“Automation”, “Markets” and “Collaborative Technologies”

“Computer automation has helped people to perform the many critical cognitive tasks that computers are unable to address, and in so doing has greatly improved their lives. Somewhat ironically, one of the most important cognitive tasks facing people is establishing the proper balance of computer automation and manual involvement. This need for balance is exemplified in the creation and operation of e-marketplaces, in which goods and services are bought and sold in high volumes across electronic networks”[Langabeer, 2001]

The four main phrases associated with ‘Collaborative Technologies’ are ‘e-commerce’, ‘content’, ‘communication’, and ‘customers’. E-commerce includes different types of sourcing methods that are used to automate e-marketplace transactions. Therefore it forms the baseline process of doing business. Customer needs could be understood by collecting customer information and personalizing it through content management. This can be done using click-stream, profile and purchase-history information.

FOUR MAIN PHRASES OF COLLABORATIVE TECHNOLOGIES

- E-Commerce
- Content
- Communication
- Customers

“The key strategies of successful e-marketplaces include mapping external data to a common internal format, standardizing on the processes for content management, phasing in different content sources over time, disturbing tools to suppliers so they can manage portions of their own content and incorporating third-party content aggregators to provide rationalized data”.[Hilderth, 2000]

“Process”, “Investments” and “Customers”

The relationship between these three keywords is meaningful in the context of business process improvement of systems having significance with customers, and thereby increasing the return on investments. The business processes are improved for building efficient systems. These efficient systems thereby increase customer value. Customer Relationship Management (CRM) technologies are used for improving relationship with customers. There is also a growing need for the organizations to justify investments in these poor economical conditions. For this justification companies look for increase in overall Return on Investments (ROI) from e-business through CRM solutions. On going through factor maps, co-occurrence maps, and white papers, it is evident that clients want several key items from their CRM vendors and integrators. The key items are such as shorter project phases, earlier and more visible results, help in thinking through business processes, help quantify ROI, metrics for CRM measurement, organizational changes through CRM initiatives, and expertise in data and process integration technologies.

“Enterprise Integration Strategy”

The ‘programs and documentations’ as well as ‘databases and files’ are two classes of legacy system assets. The former could be classified as ‘Business process assets’ and the latter, ‘Information assets’. A complete IT system integration initiative should involve both these assets in proper combination.

From the white papers, the complimentary classes of infrastructure software that addresses the integration of legacy system assets are ‘Application layer integration’ and ‘Data layer integration’. The types of system integration are already explained in the introductory chapters. The diversity of the IT assets of midsize to large companies will require a combination of business and information assets. This

combination would be provided by software infrastructure products. [Cadarette, Durward 2001]

“Application Integration”

This approach means that the applications contain the business logic of the enterprise. The business logic has to be preserved by extending the application interfaces to interoperate with one another.

There are four classes of application layer integration solutions. They are “Screen scrapers”, “Ad hoc connectors”, “Centralized hub and message routers”, and “integration brokers”.

“Data Integration”

The philosophy of data integration is to gain direct values to the enterprise through the business logic in the data and metadata which can be easily manipulated by applications in the new architecture of the enterprise. Information integration within the latest applications has gained popularity in the latest business application scenario. Although single database infrastructure is ideal for all applications within a business, it is unattainable due to evolutionary nature of IT investments. The four classes of data integration are “staged integration”, “replication”, “federated databases”, and “data integration servers”.

“Java Technologies” and “Legacy Systems Integration”

Web enablement of applications has become a growing need to support the integration of new technologies. In legacy systems, point-to-point integration existed for integration. A wide range of old and new custom and customer-off-the shelf (COTS) software systems has to be found for Java and server technologies for integration with legacy systems. [Scala, 2000]. For example, Building a J2EE based hub and spoke architecture using BEA-Web logic integration as the hub is required for this kind of integration

with the legacy systems. This includes a combined application, business process management, and integration server. Sizing the server to meet the needs of the application has been increasingly allowed by Linux.

“XML Technologies”, “IBM”, “Microsoft” and “Customers”

XML interfaces are common for products in every application software vendor in every software industry category. The major software developers in this area are IBM and Microsoft. For example, Microsoft has developed new frameworks, tools, protocols and development environments that exploit XML features to accelerate the construction of XML based applications. XML parsers and XML stylesheet transformers are readily available under public licenses. “There are over 60 industry working groups that bring technology professionals from competitors together with vendors to study and define industry-specific message conversions from EDI to XML, new messages and the higher order issue of business process.” In short XML services have become cornerstone just like Web Services.

THESIS STATEMENT

The interpreted text mined information is further used for finding answers for research questions. The selected white papers are read in depth for finding out key usage and benefits of white papers. The thesis statement is as follows,

“Industry white papers are used for better decision making by managers through better understanding of latest technologies, because the main factor involved in this process is effective communication by technologists to make managers understand technologies for business in a timely manner.”

BEST PRACTICES FOR INFORMATION SYSTEM INTEGRATION

Best Practice No. 1
Building systems with deployment in mind: System integration is not real unless a common solution is deployed to the entire organization. But normally in early stages of any system integration project, the focus is toward technology selection. This is due to the fact that technology comparisons could be complex and time-consuming for the project team. Despite this for best system integration results it is vital to keep the end goal in mind. [Yee, 2002]

Best Practice No. 2
Early and Ongoing Performance Profiles: The failure of many system integration projects happen due to performance issues at the time of deployment. This can be rectified by 'Continuous Performance Profiling' during the process of development.

Best Practice No. 3
Build a Traceable System: Tracing code allows tracking the progress of data and execution of code segments as the data is processed. Also detecting problems in distributed integration architectures is tricky. Hence the system should be instrumented with tracing code.

Best Practice No. 4
Reviewing and Addressing Secondary Scenario: Secondary scenarios arise when the data flow is not completed with expected course of execution. An example for integrated flow is illustrated in the steps as follows. Suppose, a customer searches for a product online. He enters the product data into the corporate website. The data is used to look up for more information from a product catalog database. Rules combining product catalog data and customer information are applied. An entry is made into CRM application about informing the sales staff of a customer's status. When such integrated flow patterns

are not followed secondary scenarios arise. Such scenarios should be avoided as far as possible.

RESEARCH ANSWERS

The support for thesis statements and answers for research questions are provided in the following pages in the form of research answers.

SYSTEM INTEGRATION BUSINESS MODELS

Introduction to Business Models

There are various definitions available for Business Models. A few of the published definitions which are popular are as follows (VASSILOPOULOU, POULOUDI, PATRONIDOU and POULYMENAKOU, “E-Business Models: A Proposed Framework, 2001)

DEFINITIONS

“An architecture for the product, service and information flows, including a description of the various business actors and their roles; and a description of the potential benefits for the various business actors; and the descriptions of sources of revenues.” [Timmers, 1998]

“An architecture of a firm and its network of partners for creating, marketing and delivering value and relationship capital to one or several segment of customers in order to generate profitable and sustainable revenue streams.” [Pigneur, 2000]

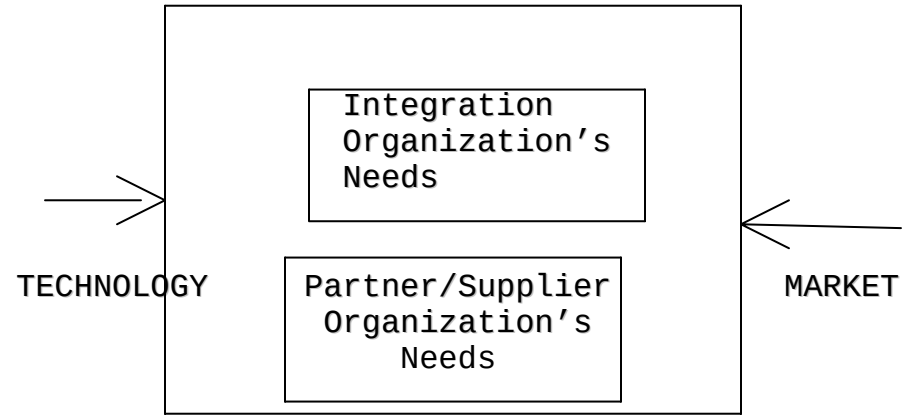
“A business model is about the invention of new value propositions that transform the rules of competition, and mobilize people and resources to unprecedented levels of performance.” [Tapscott, 2000]

“A description of the commercial relationship between a business enterprise and the products and/or services it provides in the market. More specifically, it is a way of structuring various cost and revenue streams such that a business becomes viable, usually in the sense of being able to sustain itself on the basis of the income it generates.” [Hawkins, 2001]

Existing Roadmap for Business Models

The roadmap exists for e-business models. The construction of roadmap for e-business models have been done by 'Eriksson' and 'Magnus' in 2000. [Eriksson, 2000] The generic business model building rule is to construct a framework in order to create a more concrete and stable environment.

Figure 5.a Framework for System Integration Business Model



Roadmap for System Integration Business Models

The system integration business model environment is guided by basically three forces, existing technology of the firm, existing technology of the partner or integrating firm, market environment. Modifying Metheli's generic business model for systems integration, the components of business model are 'Direction of integration', 'Integration Strategy', 'Control Model', 'Service Integration', 'and Forms of Co-operation'. [Methlie, 2001]

Analysis Framework for System Integration Models

The themes for adoption of e-business models are classified as organizational, societal, individual and technical. [Vassilpolou, Pouloudi, Patronidou,

Poulymenokou, 2001] This paper extends the same themes for systems integration business models.

Organizational: Successful 'System Integration' initiatives will lead to change in nature and structure of work in inter and intra-enterprises. Such an impact on the organization will create new communities among partners, customers, and competitors.

Societal: System integration business models will have its own industry best practices just as any other business models possess. Some of the 'Best Practices' are 'Building with deployment in mind', 'Early and Ongoing Performance Profiles', 'Building a Traceable System', 'Reviewing and Addressing Secondary Scenarios'.

Individual: The users of the integrated system such as employees, customers and partners are given more importance due to user acceptance concerns. The final integrated system could create an impact on their decision making process.

Technical: The integration technology advancement will extend the value-chain beyond with more interaction between partners, suppliers, and customers. This may in turn gain success and market penetration in different systems integration technology areas such as XML, WASP .etc. just as e-business did a few years ago. The identification of impact of specific technologies in different intra and inter-enterprise system environment would be more useful for adoption of a specific business model for the growth of organizations.

Business Performance Improvement

[Dutta & Manzoni, 1999] developed a framework on business performance improvement for evaluation of specific contribution of technology or market change. According to this research there are two forces affecting the market structure. They are 'Business Opportunities' and 'Technology Advances'. The relation

between these two forces is plotted as graph in four quadrants. [Dutta, Manzoni, 1999]

Quadrant-1: Remarkable changes are not noticed.

Quadrant-2: Technological advances are much higher and that will affect the way of cooperation. The cooperation in this context means the network between the firm, its suppliers, partners and customers. Due to the rapid advancement of integrated system there would be an affect in these networks.

Quadrant-3: The amount of value in this quadrant is defined by new business opportunities due to system integration technologies. The visibility of growth in business opportunities would decrease the barrier entries of specific industries in the value-chain. This means that when system integration technology increases the business growth in a specific industry, the reason for which could not be traced easily. This might cause a barrier to other participants in the value-chain.

Quadrant-4: This is an ideal case where high competition force all the firms and industries to be organized in an effective and efficient manner. Whether or not a firm will participate in the value-chain is decided by product differentiation and service differentiation across industries. [Eric, 2000]

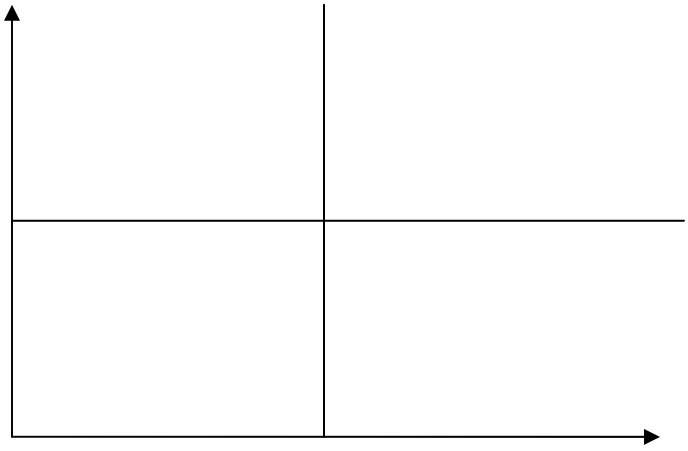
By placing the project situation on the graph, the improvement in business performance after system integration implementation could be identified.

Higher values could be gained by technology push, investing in new system integration technologies or markets pull, getting a better business opportunity [Morton, 1991]. Scott Morton proposed that it is better to examine one dimension (technology or market) first and later move on to another. Technology push is the method in which the advancement of system integration is done before business opportunity achievements. In market pull higher values through business opportunity achievements are examined before the adoption of system integration technologies.

Figure 5.b: Factors influencing System Integration business model

(Y-AXIS) BUSINESS OPPORTUNITIES

HIGH Q3 Q4



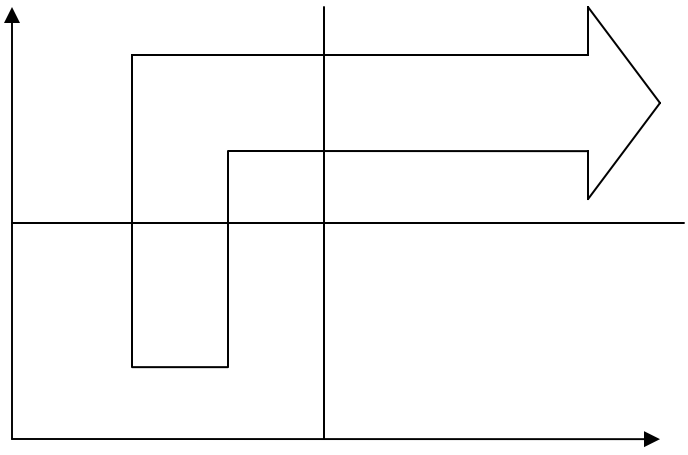
LOW Q1 Q2 HIGH

(X-AXIS) SI TECHNOLOGY ADVANCEMENT

Figure 5.c: MARKET PULL

(Y-AXIS) BUSINESS OPPORTUNITIES

HIGH Q3 Q4

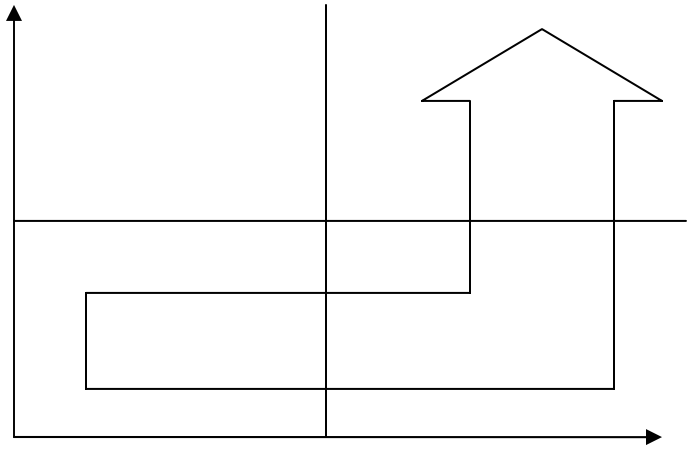


LOW Q1 Q2

(X-AXIS) SI TECHNOLOGY ADVANCEMENT

Figure 5.d: TECHNOLOGY PUSH
(Y-AXIS) BUSINESS OPPORTUNITIES

HIGH Q3 Q4



LOW Q1 Q2

(X-AXIS) SI TECHNOLOGY ADVANCEMENT

Assessment of System Integration Business Models

Assessment and appraisal of business models are done by Aftah and Tucci (2001). This model defines three levels for measuring business model performance [Pateli, 2001]. The factors for evaluating a generic business model are extended to system integration business models.

a) Profitability Measures: In a generic business model, comparison of a firm's profitability is done to that of competitors or between departments in an organization using measures such as cash flows and earnings. In system integration business, such comparisons cannot be made. Rather the profitability should lead to the whole enterprise in case of intra-enterprise integration and to all integrating firms in case of inter-enterprise integration.

b) Profitability Prediction: This concerns firm's profit margins, revenue growth rate and revenue market share with those of industry competitors. These measures rely on components of business model and how those components are related to each other.

c) Business Model Component Attributes: In order to assess the capabilities of each component of the business model, system integration firms should themselves 'To what extent are the firm's capabilities distinctive, inimitable, and extendable to other product markets?". A high rate is given to the component if the answer to the question is 'yes'.

METHODS FOR CALCULATING THE ROI OF SYSTEM INTEGRATION PROJECTS

The traditional thinking of Information Technology (IT) as a cost center has re-appeared in most enterprises. IT-budget cutbacks have caused a re-emphasis of Return on Investments (ROI) calculations. ROI computation for system integration projects is challenging due to their complexity and high risk. While intra-enterprise systems integration projects are risky, there is an even greater need to justify expenses when more than one enterprise is involved in the integration process. Calculation of ROI for inter-enterprise integration is more challenging due to due to the lack of availability of data, different and varying degrees of need for ROI based justification, and different internal values of the integrating enterprises. Many of the returns that organizations expect from successful inter-enterprise integration, such as decreased time to market for new products, enhanced customer satisfaction, technological excellence, and market leadership are intangible. These projects aim to achieve seamless integration of disparate systems existing on separate platforms managed by diverse departments, customers, suppliers, and other partners. ROI calculations for such projects are based on numerous intangible factors for both returns and risks. This does not mean that ROI should not be computed for these projects. On the contrary, it means that a better roadmap needs to be used to develop a robust ROI measure.

ROI for ‘System Integration’ in ‘Mergers and Acquisitions’ already exist, such as the one which was developed by National Institute of Science and Technology, US Department of Commerce. But roadmap for ROI of system integration does not exist. This paper suggests a roadmap for ROI of system integration.

Existing ROI for System Integration

A tactical roadmap to system integration in mergers and acquisitions exist. It is followed in four phases. The first phase assists in the development of actionable guiding principles. The current state of the IT infrastructure at both the integrating companies would be assessed. The risks and opportunities would be found out by knowing the goals for the integrated system and business processes. The second phase illustrates the optimized main business processes and merged organization roles. Data analysis profile is created in the third phase. The main data integration issues to be addressed in order to satisfy internal processes for the merged organization would be addressed. The steps for generating files for interaction with customers, partners, and suppliers are laid out. In the last phase, a detailed inventory of the technologies at the merging companies will be completed. A survey is taken to evaluate the achievement of business objectives of the merged organization.

A similar roadmap for the calculation of ROI for systems integration projects is illustrated in this paper.

Business Imperatives

The business benefits of the systems integration projects are:

1) Increased direct labor, material, management, and customer productivity through seamless integration
2) Revenues gains from new modes of business that are possible due to integration
3) Strategic goals achieved due to integration

Tangible ROI measures can be appraised at an actual or approximate value, whereas intangible ROI

cannot be measured in terms of a numerical figure. Intangible measure include subjective methods of appraisal such as increased usage of sophisticated technologies, improvement in processes, customer satisfaction, market share changes, Quality consciousness, better strategic planning and decision making, human resource achievements.

Phases in ROI Calculation (Roadmap)

The following steps could be followed to calculate ROI for any systems integration project:

1) Standards Based Integration: Most system integration projects focus on solutions to immediate problems in linking data and processes. As a result, ROI calculations often focus on short term costs and benefits. However, such small integration projects form a patchwork of linkages between systems that need extensive maintenance in the long run. ROI computation with a narrow focus is valid only for short-term projects and turn out to be worthless in estimating the long term benefits of the projects. Businesses should determine the dominant strategy for information system integration in the industry. For a better ROI of long-term system integration projects, the standards that are currently dominant in the industry, or that are likely to dominate in the future have to be followed.

2) Process Maps: In any inter-enterprise systems integration project, the process maps of both the integrating company and partnering companies have to well understood, documented, and the systems should be updated. A business process map describes the chronological sequence of event functions and includes a description of data, organization, and resources that are directly associated with the event functions [Scheer, Abholhassen, Jost, Kirchmer Business Process Excellence, p16 2002]. If there is any flaw in one of the partnering organization’s process maps, then there are higher chances of failures of system integration. There is a possibility of doing a backward check in case of updated systems. In the same way, for intra-enterprise integration, the process maps of the organization have to be thoroughly understood and

documented. Good ROI can be found out only when process maps are well-known.

3) Employee/User Skill Development: The users or employees who use the integrated system have to be provided end-user education. The educated employees could transform the organization into a knowledge-based organization. Also retrenching has to be reduced through better training for employees which will improve the knowledge-base and skill-sets of the organization who work collaboratively. ROI calculations for integration projects are meaningful only when employee grounds are laid out.

4) Technical Solution: It is necessary to have a broader picture of the enterprise in mind when writing the plan for a system integration project. Such initiatives should be considered as a business need. There should be an improved process and better workflow before the start of integration. Effective execution of integration plans have to be ensured which will provide systems adaptable to ecosystems.

5) ROI Calculation: The ROI should be calculated only when the above four steps are complete. In case of intra-enterprise integration, both tangible and intangible ROI could be found out using one or more methods as described in this paper. For inter-enterprise integration, it is hard to find out tangible ROI when there is latency between the integrating organizations. In such cases, only “intangible ROI” could be assessed.

Calculation of Tangible ROI

The formula for calculating the return of investment is,
Return on Investments, ROI = (C-NPV)/C where, C=Cost of Investment

NPV= Net Present Value of Savings

Also, financial formulae or models the return on investments (ROI), Internal Rate of Return (IRR), Net Present Value (NPV) and Payback period could be calculated.

Cost-Benefit Analysis is the simplest way of measuring whether the returns are positive or not. It is measured by simply subtracting benefits from costs.

Economic Value-Added, EVA is a registered trademark of Stern Stewart & Co.; EVA is a method of periodic financial value-based performance measure. It focuses on maximizing the value of shareholders (shareholder's wealth), by calculating the true economic profit of the company. Hence, it helps managers to create value to shareholders.

EVA = OPBT – Tax – (TCE*CC)

Where, OPBT=Operating profit before taxes
TCE=Total Capital Employed
Taxes=Federal and State taxes
CC=Cost of Capital

Balanced Scorecard methodology is an analysis technique designed to translate an organization's mission statement and overall business strategy into specific, quantifiable goals and to monitor the organization's performance in terms of achieving these goals. Hence it could be classified as both tangible and intangible ROI identifier. Its methodology examines performance in four areas: financial analysis, performance indicator, customer analysis, and internal analysis. [Kaplan, Norton; The Balanced Scorecard 1996]

Value chain analysis
Profitability is measured by the long-term ROI of the system integration project, which depends on five factors. They are bargaining power of buyers, bargaining power of suppliers, and threat of substitute products or services, threat of new entrants and intensity of competition.

IT-Centrix provides a method to calculate the IT contribution to employee productivity. According to which, productivity is output produced per unit of input.

$$\textit{Productivity/employee} = \frac{\text{Revenue per employee}}{\text{Fully loaded salary}}$$

When the same method is extended to system integrated organization, the productivity per employee due to system integrated-IT contribution could be measured using the same formula before and after the system integration. The return on investment in this case is indicated by the increase or decrease in productivity per employee, which means the overall productivity of the organization is improved or decreased by this project.

Statistical tools using simulation or real-options approaches hold forth the promise of better results. However, since they need extensive inputs for their computations, they are only as reliable as the data input to the project. This makes them suitable only for System Integration projects that are small, use established technologies, and affect well-understood processes.

Intangible ROI Identification

This method is more subjective, and the results may differ from company to company or even from project to project in the same organization. System integration projects are generally done to achieve the long-term goals of the enterprise. But in reality, many companies cut costs such as investment, and labor at the beginning of the project having short-term business priorities in mind leading to failures and financial losses. Long-term strategies should be articulated by the senior management before such huge initiatives which would provide confidence to the entire project team. The project should focus on achieving that plan.
According to an AMS survey, only 25% of European integration companies have moved beyond using middleware tools to connect applications and among them 5% have been successful in extending their businesses with their partners and suppliers. This is due to integration immaturity.

Increased usage of sophisticated technologies

System integration is becoming pervasive according to many recent research results. Mercator predicts that integration technology would be applied

to larger and more varied problems, as commercial integrations mature and organizations gain experience with initial projects. This could lead to an induction of modern technologies into every type of organization. Every system integration project should ensure whether it is creating an impact in the organization. There are no proper metrics to measure this impact. A simple way to measure such an impact (returns) is by taking a qualitative or even a quantitative survey from the employees, managers, suppliers, partners, and customers after the project implementation.

Improvement in Process

When an application has to be integrated with another application from outside the organization, there would be a change in process in order to integrate them. Many vendors like Webmethods, Vitria, Iona, and Extricity are trying to provide such solutions to systems integration companies. This change in process should lead to improvements. The intangible ROI from process improvements would be employee efficiencies, reduced procurement costs, reduced inventory, increased communications efficiency, effective communication between two groups, and better workflows.

Higher Customer Satisfaction

Customer satisfaction is a primary goal of business, and nowhere is the need greater than in managing the customer experience with complex, highly integrated system. Integrated system brings together business processes from diverse organizations or departments within organizations. Customers need to perceive a seamless experience with these processes when they are integrated using information technology. While highly integrated system offer customers a single interface to multiple processes, they can become very unstable in the presence of even minor errors or high load. System integration projects need to seriously consider the impact of these new processes, and use effective monitoring to ensure that customer satisfaction levels are maintained.

Increased Market Share

Integration is an enterprise-level project. The old perception of system integration as a smaller series of patchworks and projects should change into enterprise level long-term projects. System integration projects are started with a focus of transforming the organization into a new level. During this process, care should be taken by the company for the minimal impact on market shares. Significant reduction in market shares will have a negative impact on the organization's financial health. Intangibility through market shares is measured by counting the achievement of systems integration objective versus change in market shares. The greater the number of achievements and increase in market share, more successful the intangible returns are.

Better Strategic Positioning and Decision Making

A regular check of integration costs and both tangible and intangible benefits have to be recorded. The difference has to be plotted in a graph to ensure whether there is a continuous improvement in the integration initiative. In cases of serious downfall, the plans have to be sorted out.

Quality Consciousness

Technology and business standards have to be maintained by the integrated system. In the days of many integration failures, there is always the need for quality integrated system. In order to receive better returns, a project should use technological standards such as RosettaNet, UDDI for web services .etc.

"RosettaNet is a consortium of major Information Technology, Electronic components, Semiconductor manufacturing and Telecommunications companies working to create and implement industry-wide, open e-business process standards".

"The Universal Description, Discovery and Integration (UDDI) protocol is one of the major building blocks required for successful Web services. UDDI creates a standard interoperable platform that enables companies and applications to quickly, easily,

and dynamically find and use Web services over the Internet”.

Surveys

Qualitative and quantitative polls have to be taken from team members, customers, partners, suppliers, and managers on a regular basis in order to ensure whether the project is proceeding in the right direction. The survey should include structured and unstructured formats that allow participants the opportunity to provide their own opinion and other comments and observations.

Human resource management

Integration is a complex process, and it needs dedicated personnel on the project for a longer period. On the contrary, most companies hire temporary or contract basis professionals for these projects. The project team members should be created to have a mental bonding with the organization in order to deal with such a complexity. Communication is the key to such a bonding. This could provide confidence to improve productivity as well as their expertise in dealing with critical situations. The project team members have to be with the company till the end of implementation of such systems. They need not be kept together post-implementation because quality of the implemented system will be assessed before the system is in use. Once the quality and testing team signs off the project for use the team may split.

Hi-Tech Labor

Last but not least, a system integration initiative can develop clerks into knowledge workers. The project implementation transforms the enterprise into technologically sophisticated. The normal manual processes would be converted into integrated automation. Application, interfaces, processes, objects, and data talk with each other. Knowledge workers would use such systems for day to day activities such as automated transactions through ERP systems, paperless office, and knowledge based interfaces.

MANAGING ENTERPRISE SYSTEM INTEGRATION IMPLEMENTATIONS

Enterprise Resources Planning or ERP tools have become the foundation of business intelligence for managing accounting, manufacturing, distribution and human resources of an organization. The motivation for a company to adopt ERP comes from the expectations of company growth and achieving competitive advantage through the use of technology. The building blocks to a successful ERP implementation are defining the requirements, developing a plan and implementing them. The real problem lies in implementing them. Most of the organizations fail during implementation.

The implementation phase starts with forming the right structure of an organization, allocating the resources to the teams appropriately. The steps in the implementation process are first the top-level management should meet and make executive decisions for the project. The project manager is responsible for day-to-day details and acts as an interface between the top-level management and the low-level workers. The project manager has to take care of the next steps in the ERP implementation and look for guidance for the crisis management on a consistent basis. Ensuring consensus among the project members and streamlining the process will move the project forward as planned. The implementation is critical and most companies fail to do it because, people do not understand the new system properly and have unrealistic expectations. They try to be flexible according to the package.

Some of the successful companies adopting ERP packages were JD Edwards, Oracle, and People Soft. They reaped the benefits from the decisions made over the last few years to build technology that leverages their ERP systems in their applications. SAP has made a good customer base through their e-commerce portal, mySAP.com.

One of the most noted ERP failures was FoxMeyer Drugs' Bankruptcy. After careful market research and product evaluation they purchased SAP R/3 with the goal of improving efficiency through technology. But they failed to implement it. SAP perceives it as the

management failure and blames that they had unrealistic expectations.

Market statistics reveal that US ERP sales grew from under $1billion in 1993 to $8billion in 1998 (Data quest, $14.8 billion worldwide). In 1998, (bring this up to date) US companies spent $80billion on ERP systems integration. Industry analysts expect an average growth rate of 37% per year for the next 5 years. (Dataquest, 1998).

Integration-Success or Failure?

From the above cases, it is very difficult to conclude whether integration is a success or failure. At last what makes a company success is that the top management of the company has to make realistic decisions and have gumptions to stick on to it. They should motivate people and educate people about the new technology and manage change. There lies the positive result.

The essence of any business is for the welfare of the society and making profits. The expectations from ERP is increased ROI through improvement in the business process, and savings in cost and time. This increase in ROI is a continuous process, meaning that it is not only expected during implementation but also forever. For this a right structure of the organization is essential. Generalizing, integration as a success or failure will only lead to confusion. The degree of success depends on how well an organization structures its' resources, make decisions, and motivates people before implementing it.

Reasons for Implementation Failure

The major reasons for integration implementation failure are:

1) Inadequate requirements definition
2) Poor ERP Package selection
3) Inadequate resources
4) Resistance to change
5) Code on Demand

Poor Package Selection: This occurs due to the company having inadequately defined functional requirements. It also occurs when the staff members assigned to ERP projects does not take the time to run the screens of the new system to find the software package that is adequate for their needs.

Inadequate Requirements definition*:* In most cases, this occurs due to the implementation of the same ERP as in the previous job held, without defining the requirements for the current needs. Hence, the results would be ERP package would not fit the organizational needs or the package selected takes longer to process daily tasks.

Inadequate Resources*:* Many companies attempt to save money by over-utilizing the resources like doing everything on an overtime basis.

Resistance to change*:* The lack of change management can prevent the implementation from succeeding. Resistance to change is caused by a failure to build a case for change, lack of involvement by those responsible for working with the changed processes, inadequate communication, lack of visible top management support and commitment, and arrogance.

Also the ERP implementation is costly and complex. The time taken to implement is always more than that of planned, statistically.

Code on Demand: AMR Research survey says, “Frequently, the [enterprise application integration strategy] approach to create integration code on demand, leaving IT with a mass of class spaghetti code that is extremely expensive to maintain”. [AMR Research response to a research inquiry. June 22, 2001, Contivo, Introducing Enterprise Integration Modeling, 2001] Therefore the level of difficulty in functional integration gets multiplied at the time of implementation. Even though a uniform architecture for enterprise is planned, the code required to transfer data would create a problem when every department continue to develop their own characteristics. The maintenance of functionality of the integrated system will become complicated.

Different inter-organizational views: In case of inert-organizational enterprise integrations, when both organizations do not have the same perspective of integration there might be a conflict. The difference in integration philosophy between organizations can lead to failure in goal achievements.

EFFECT OF WHITE PAPERS IN SYSTEM INTEGRATION MARKET

White papers serve as a communication channel between technology innovators (scientists/engineers) and technology introducers (managers). In modern technology era where everyday a new jargon is born, white paper helps managers understand the story behind that jargon before a viable business is created. This explains the need for better understanding of technological ideas by non-technical executives for decision making of an organization.

Ingredients of a good white paper

On going through white papers of system integration for decision making the essential ingredients of a good white paper are found. They are mentioned as follows,

1) **Technology as theme**: The purpose of white paper is to deliver a solution to a problem. [Stelzner, Michael ‘How to Write a White Paper’, 2003] White papers would be used by managers to adopt those solutions in real time for better results. Hence white papers should not be used to advertise the company. Rather the idea generated by the company should be made popular through this document.

2) **Organization of white paper**: The white paper should be organized in a sequential manner such as introduction, business needs, practical problems, technology solutions .etc. This will provide a better structure for the writer to think and communicate to the reader, as well as the reader to verify quickly when he or she runs short of time. Hence better structure of the paper would allow the reader to grasp more ideas.

3) **Method of communication**: The communication approach should be to convince the decision maker to adopt the intended solution to gain benefits. There are two ways to write white papers: (a) focus on self-interests (b) concentrate on the reader's interests. [Stelzner, 2003] The former method concentrates more on product thereby makes the reader get bored. Whereas the latter is a customer focused approach. This method explores customer's problems and suggests solutions for those problems.

4) **Cite Examples**: Suitable examples have to be cited in order to make the reader get involved through practical scenarios. The examples can stimulate the creativity of readers to utilize ideas for different problems.

5) **Content**: The content of the white paper should be original. References, if any should be cited wherever needed. The language should not be verbose and the ideas should be precise to the problems. The minimum number of ideas should be supported by maximum evidence in order for the content to be more valuable.

6) **Size**: The normal length of the white paper according to Stelzner is 11 pages. But this number can be varied depending on the topic and purpose of the white paper.

7) **Additional information**: The additional information such as diagrams, charts, pictures, graphs .etc. should be suitably illustrated wherever it is needed. This will help the reader to understand the data clearly and interpret more details from the white paper.

8) **Outlook**: The outlook of white papers should be pleasing to the reader. Readability formulas could be used to verify the outlook of white paper. Gaudy colors should be avoided. The font type and font size should be uniform throughout the document. Better readability helps readers recognize information easily.

Recommendations for better results to readers through white papers

Writer of White Papers: The selection of white paper writer is critical because he or she is the person who will present the ideas to the readers who have some or little knowledge about the subject of the white paper. The writer should possess technical knowledge of the relevant subject. This is because unless the writer does not have the grasp of the subject, he or she could not critically analyze and present his or her ideas in a coherent manner. This does not mean that the writer should be the person who develops the product or ideas, but he should be someone who clearly understands the technical aspects as well as the intersection of it with business needs. Hence in normal cases writer could not be someone from marketing department or an independent consultant. In general engineers or technical professional possess less proficiency in writing or time constraints due to technical duties. Hence the writer should suitably be a senior technical executive who possess significant professional experience, and superior technical leadership and skills in the industry.

Better interpretations for better decisions: The main users of white papers are managers who use it as a source of updating their knowledge and trends about different technologies in the market. They read about emerging technologies and make strategic and tactical plans for using technologies for the improvement of business. Hence the writer of white paper should emphasize more about technologies in such a manner in which managers would understand. A completely technical paper would not be understood by everyone. Managers have to be provided information in such a manner that they would interpret the technology’s intersection with business needs, and use the idea for their decision making.

Stay ahead in market through technology sophistication: White papers are used to convey technology information to everyone in order to maintain knowledge integrity in the organization as well as in the market. Its goal will be materialized

only when it is implemented in real time. When the implementation takes place, it directly means that the idea has been put into practice. Hence white paper has been successful in spreading ideas to decision makers. This will help implementing organization differentiate itself in the market through early adoption of new technologies or technology sophistication.

Chapter 6
Discussion

INDIRECT FINDINGS THROUGH QUALITATIVE ANALYSIS

The indirect findings in this research are completely qualitative. This involved reading of white papers in the area of 'Information System Integration' after mining the unstructured texts.

The list of indirect findings is as follows,

1) For better Return on Investment (ROI) calculations of system integration projects, both tangible and intangible benefits (short-term and long-term) should be taken into account.
2) The factors for evaluation of generic business models could be extended to system integration business models. They are profitability measures, profitability prediction, and business model component attributes.
3) The major reasons for system integration implementation failures are poor package selection, inadequate requirements definition, inadequate resources, resistance to change, code-on-demand, and difference in inter-organizational philosophies.

DIRECT FINDINGS THROUGH THEORETICAL ANALYSIS

The direct findings of this research are based on theoretical analysis as well as based on the experience of the study of industry white papers.

The list of direct findings is as follows,

1) The writer of white papers should ideally be a senior technical executive who has significant professional experience and superior technical skills.
2) Managers should be presented white papers that would help them to interpret technology's

intersection with business needs. Such a presentation would help managers make better strategic planning.

3) Adopt white paper ideas in real time to realize successful results. The implementing organization would differentiate itself by staying ahead through technology sophistication.

VALUE OF TEXT MINING IN THIS RESEARCH

The value of text mining is realized at the end of the research when surprising results are seen. Some of the surprising results were “XML” and “Application Servers”. XML is common for products in application software. When this relationship is extended further, it is seen that Microsoft which produces application software has developed new frameworks, tools, protocols, and environments that exploit XML features to accelerate the construction of XML based features. A different relationship that was seen was between the words, *strategy* and *outsourcing.* The relationship was depicted only after combining the words together as *strategic outsourcing*. Hence text mining was a powerful tool for identification of such relationships.

POTENTIAL PAPERS FOR PUBLICATION FROM THIS THESES

1) “Return on Investments of System Integration Projects”
2) “System Integration Business Models”
3) “Managing System Integration Implementation Failures”
4) “Recommendations for Better ERP Implementations”
5) “White papers improve decision making in IT enterprises”
6) “Effect of timing in the introduction of technological details to Corporate-IT managers”

RECOMMENDATIONS FOR FURTHER RESEARCH

The following listed research could be some of the potential future works that could be pursued.

1) Research in a particular information systems integration area, such as Data Integration, Application Integration, and Process Integration.
2) Different method of usage of text mining tool for research in systems integration.
3) Research in Inter-organizational Process changes in system integrated organizations.
4) Further development of ROI calculations of System Integration projects.
5) A survey of popular system integration business models currently used in the industry.
6) A study of managerial decision-making methods through whitepaper reading in technology organizations.

BIBLIOGRAPHY

[Synergy Connex] Synergy Connex, “Integration provides a seamless link between all business processes – Software components across the value chain”, SCX-0301-01, (2001).

[Leem, Seong et al. 2002] Shin, Kitae; Leem, Choon Seong, “A reference system for internet based inter-enterprise electronic commerce”, The Journal of Systems and Software 60 (2002) 195-209, (March 2001).

[DePryker et al.1993] DePryker, M., Peschi, R., and Landegem, T., “B-ISDN and the OSI Protocol Reference Model,” IEEE Network, Vol 7, No. 2, pp.10-18, New York, NY (March 1993).

[Satyanarayanan 1990a] Satyanarayanan, M., “A Survey of Distributed File Systems,” Annual Review of Computer Science, Vol. 4, pp. 73-104 (1990).

[Sandberg 1987] Sandberg, R., “The Sun Network File System: Design, Implementation and Experience,” In: Proceedings of the USENIX Summer Conference, USENIX Association, Berkeley, CA, pp. 300-314 (June 1987).

[Gray 1978] Gray, J.N., “Notes on Database Operating Systems,” Lecture Notes in Computer Science, Vol. 60, Springer-Verlag, Berlin, pp. 393-481 (1978).

Bisbal, J., Lawless, D., Wu, B., and Grimson, J, Legacy Information Systems: Issues and Directions, IEEE Software, 16(5), Sep/Oct 1999, 103-111

Cadareete, Durward, “Achieving a Complete Enterprise Integration Strategy”, December 2001

Chalmeta, R., Campos, C., Grangel, R, References Architectures for enterprise integration, Journal of Systems and Software, 57(3), 2001, 175-191

Chatfield, A.T. and Yetton, P. Strategic Payoff from EDI as a Function of EDI Embeddedness, Journal of MIS, 16(4), Spring 2000, 195-224

Clark, T.H and Stoddard, D.B. Interorganizational business process redesign: merging technological and process innovation, Journal of MIS, 13(2), 1996, 9-28

Dalrymple, Kelly “TCO Analysis: The Bottom line of Outsourcing ERP”, Blue Star Solutions, September 2002

Dutta S, Manzoni J, (1999) “Process Re-engineering, Organizational Change and Performance Improvement”, McGraw Hill, 1999.

Eric S, “Building E-Business Models: An Analytical Framework and Developmental Guidelines”, 2000.

Eriksson H, Magnus P, (2000) “Business Modeling with UML-Business Patterns at Work”, John Wiley and Sons Inc, 2000.

Fayad, M.E., Hamu, D.S., and Brugali, D. Enterprise Frameworks characteristics, criteria, and challenges”, Communications of the ACM, 43(10), Oct. 2000, 39-46

Flurry, G and Vicknair, W. IBM Application Framework for e-business, IBM Systems Journal, 40(1), 2001

Hargis, Gretchen, IBM Corp., Readability and computer documentation, ACM Journal of Computer Documentation, August 2000/Vol.24, No.3

Hasselbring, Information Systems Integration, Communications of the ACM, 43(6), June 2000, 32-38

Hawkins R. The “Business Model” as a Research Problem in Electronic Commerce. Issue Report N.4.2001.

Hilderth, Sue “From e-Markets to e-Networks”, ebizq, October 2000

Jeusfeld, A. Manfred, Moor, de Aldo, “Concept Integration Precedes Enterprise Integration”, Proceedings of the 34th Hawaii International Conference on System Sciences-2001

Klare, R. George, The measurement of Readability: Useful Information for Communicators, ACM Journal of Computer Documentation, August 2000/Vol. 24, No. 3; pp107-121

Langabeer, Jim “Automating the Demand Chain”, ebizq, May-2001

Lassila, K. and Brancheau, J. Adoption and Utilization of Commercial Software packages; Exploring Utilization Equilibria, Transitions, Triggers, and Tracks, Journal of MIS, 16(2), Fall 1999, 63-90

Leem, C.S. , Kim, S. Introduction to an integrated methodology for development and implementation of enterprise information systems, The Journal of Systems and Software, 60(3), 2002, 249-261

Malone, T. and Crowston, K. The interdisciplinary study of coordination, ACM Computing Survey, 26(1), 1994, 87-119

Mercator, Integration Technology Vision, Version-3, April-2002, pp 11-12

Mercator Software, “Integration Megatrends: The Current and Future State of the Technology”, Intelligent Business Integration Team (Jan 2003)

Methlie, L. B. (2001) “New platforms for commerce: Integration and collaboration in the electronic marketplace” (in Norwegian). In S. Ulset (Ed.). Fra summetone til informasjonsportal – SNF arbok 2001, Bergen, NO: Fagbokforlaget.

Pateli, Adamentia (2001) “Whitepaper: A Domain Area Report on Business Models” pp35-36, Athens University of Economics & Business, 2001.

Pigneur Y. The e-Business Model Handbook, HEC working paper. 2000. Available Online at: http://216.239.51.100/search?q=cache:2hVu0z1adLwC:inforge.unil.ch/yp/Pub/00-ebmh.pdf+Tapscott+D+AND+Business+Model&hl=en&ie=UTF8

Purao, S., Jain, H., and Nazareth, D. An Approach to Distribution of Object-Oriented Applications in Loosely Coupled Networks, Journal of MIS, 18(3) Winter 2001, 195-234

Sang, J., Follen, G., Kim, C., Lopez, I., Townsend, S. Migrating Scientific Applications towards CORBA-based client-server architectures, Software Practice and Experience, 31(14), 2001, 1313-1330
Scala Steven, "Meeting the Challenges of Web-Portal Integration", ebizq, November-2000

Scott-Morton M. (1991) "The Corporations of the 1990s: Information Technology and Organizational Transformation", Oxford Press, 1991.

Shin, K., Leem, C.S., A reference system for internet based inter-enterprise electronic commerce, The Journal of Systems and Software, 60(3) 2002, 195-209

Smith, D. Tom, "White Paper, searchWebservices.com definitions", Oct, 2002;
http://searchwebservices.techtarget.com/sDefinition/0,,sid26_gci213361,00.html

Spinuzzi, Clay Exploring the Blind Spot: Audience, purpose and Context in "Product, Process, and Profit" November 2000/Vol.24, No.4 ACM Journal of Computer Documentation; pp213-225

Stelzner, Michael 'How to Write a White Paper-A White Paper on White Papers', 2003

Strehlo, Kevin, Publisher, Enterprise Development Journal; Data Junction- Integrating the Interconnected World, "Application Integration: Tactics to Data Transformation Nirvana" pp4-5

Sumi, T. and Tsuruoka, M. Ramp new enterprise information systems in a merger and acquisition environment; a case study, Journal of Engineering and Technology Management, 19(1) 2002, 93-104

Tapscott D. Growing up digital: The Rise of the Net Generation. 2000

Thmistocleus Marinos, Irani Zahir, Paul Ray, Keefe Robert, “ERP Problems and Application Integration Issues: An Empirical Survey”, Proceedings of the 34th Hawaii International Conference on System Sciences - 2001

Timmers, P. Business Models for Electronic Markets, Journal on Electronic Markets, (2): 3-8. 1998.

Vassilpoulou, Pouloudi, Patronidou, Poulymenakou, “E-Business models: A Proposed Framework”, Athens University of Economics & Business, 2001.

[Watson, Schneider et al]; Watson, F. Edward, Schneider, Helmut, “Enterprise Systems and Business Integration”, Manufacturing report; pp.128-134, 2001

Yee, Andre “Best Practices for Successful EAI”, ebizq, July 2002

Zibell, Kristin, Klare’s “Useful Information” is Useful for web designers, ACM Journal of Computer Documentation August 2000, Vol.24, No. 3, pp141-147

APPENDICES

APPENDIX A: GLOSSARY DEFINITION OF INFORMATION SYSTEM INTEGRATION TERMS

API Application Program Interface - an interface that enables programs to communicate with each other.

APPC Advanced Program-to-Program Communication— IBM's solution for program-to-program communication, distributed transaction processing and remote data access across the IBM product line.

Application Service Providers Companies which allow business customers to rent or lease the use of e-business applications over secured Internet connections.

Application to Application Integration (A2A) A form of enterprise application integration in which two or more applications, usually but not exclusively within the same organization, are linked.

Asynchronous Communications A form of communication by which applications can operate independently, so that they do not have to be running or available simultaneously. A process sends a request and may or may not wait for a response. It is a non-blocking communications style. (Or) An interaction between two autonomous entities in which the initiator does not suspend its execution (i.e., block) after the communication event; it does not expect any response from the recipient.

Automatic Binding Describes the action when an RPC client stub locates a specific server on a list of servers.

B2B Business to Business. In reference to commerce conducted between companies rather than between companies and private individuals or consumers.

B2C Business to Consumer. This refers to commerce conducted between companies and individuals or consumers. In contrast to B2B.

Backbone A series of connections that forms a major communications pathway within a network.

Bandwidth The amount of data that can be sent through a connection; usually measured in bits per second. A fast modem can move about 15,000 bits in one second (about a page of English text).

Binding The association of a client and a server.

BizTalk A Microsoft sponsored set of guidelines for how to publish schemas in XML and how to use XML messages to easily integrate software programs together in order to build rich new solutions.

Blocking Communications A synchronous messaging process whereby the requestor of a service must wait until a response is received.

Buffered Queue A message queue that resides in memory.

Business Process Management The concept of shepherding work items through a multi-step process. The items are identified and tracked as they move through each step, with either specified people or applications processing the information. The process flow is determined by process logic and the applications (or processes) they play virtually no role in determining where the messages are sent.

COM Component Object Model— Microsoft's standard for distributed objects, an object encapsulation technology that specifies interfaces between component objects within a single application or between applications. It separates the interface from the implementation and provides APIs for dynamically locating objects and for loading and invoking them (see DCOM).

Communications Middleware Software that provides inter-application connectivity based on communication styles such as message queuing, ORBs and publish/subscribe.

Communications Protocol A formally defined system for controlling the exchange of information over a network or communications channel.

Connectionless Communications Communications that do not require a dedicated connection or session between applications.

CORBA Common Object Request Broker Architecture— a standard maintained by the OMG.

CPI-C Common Programming Interface-Communications— IBM's SNA peer-to-peer API that can run over SNA and TCP/IP. It masks the complexity of APPC.

Customer Relationship Management (CRM) Customer Relationship Management (CRM) is the art of integrating every aspect of IT that relates to the customer – quite literally everything from marketing through sales to accounts receivable and bill collection.

Customer Relationship Marketing (CReM) Using a company's knowledge and understanding of a customer to more effectively market to that customer. While most often used in B2C e-commerce, CReM techniques are still frequently employed in B2B situations.

cXML Commerce XML. A meta-language that defines the necessary information about a product. It is a set of document type definitions (DTD) for the XML specification. Eventually, it will be used to define the exchange of transaction data for secure electronic transaction over the Internet.

Data Level Integration A form of EAI that integrates different data stores to allow the sharing of information among applications. It requires the loading of data directly into the database via its native interface and does not involve the changing of business logic.

Data Knowledge and Decision Support Using IT technology and information to both measure and improve an organization's performance.

Data Transformation A key requirement of EAI and message brokers. There are two basic kinds: syntactic translation changes one data set into another (such as different date or number formats), while semantic transformation changes data based on the underlying data definitions or meaning.

Database Middleware Allows clients to invoke SQL-based services across multi-vendor databases. This middleware

is defined by de facto standards such as ODBC, DRDA, RDA, etc.

DCE Distributed Computing Environment— from the Open Software Foundation, DCE provides key distributed technologies such as RPC, distributed naming service, time synchronization service, and distributed file system and network security.

DCOM Distributed Component Object Model— Microsoft's protocol that enables software components to communicate directly over a network in a reliable, secure, and efficient manner. DCOM is based on the DCE-RPC specification and works with both Java applets and ActiveX components through its use of the COM.

Directory Services A way for clients to locate services. Usually contained in a single system image of available servers.

Distributed Application: An application that executes on a distributed system and in which parts of the application execute on distinct computing entities. Examples of distributed applications include an automated business process that is composed of distinct business systems and resources, and end-to-end network services that are delivered over multiple network elements.

Distributed Middleware: A collection of programmer tools (utilities, libraries, services), that shields programmers of distributed applications from the difficulties that arise from executing and programming in a distributed system. In particular, heterogeneity (of hardware and software), the need to account for remote failures, the difficulty of knowing the exact state of the distributed application, location and characteristics of hardware software and network resources. Distributed middleware may, itself, be a distributed application.

Distributed System: A collection of autonomous physically dispersed computing entities (hardware or software) that are connected by some communication medium.

DOM Document Object Model - a platform- and language-neutral interface that will allow programs and scripts to

dynamically access and update the content, structure and style of documents.

DRDA IBM’s Distributed Relational Database Architecture.

EAI Enterprise Application Integration is a set of technologies that allows the movement and exchange of information between different applications and business processes within and between organizations.

E-Business Also "E-Biz." The use of Internet technologies, and the Web in particular, to conduct business operations.

E-Procurement Using the Internet and other open networks to perform the procurement function; a B2B application.

Extranet A network that links an enterprise to its various divisions and business partners which uses secured Internet links. In this way, it has the advantages of a private network at the cost of a public one. See VPN.

Federation: Hierarchical grouping of the resources (hardware, software, network) in a distributed system. The structure of the hierarchy may be policy-based or rules-based, and policies may be changed seamlessly.

Gateway A hardware and/or software setup that performs translations between disparate protocols.

Groupware A collection of technologies that allows the representation of complex processes that center around collaborative human activities. It is a model for client/server computing based on five foundation technologies: multimedia document management, workflow, email, conferencing and scheduling.

Heterogeneity A typical enterprise information system today includes many types of computer technology, from PCs to mainframes. These include a wide variety of different operating systems, application software and in-house developed applications. EAI solves the complex problem of making a heterogeneous infrastructure more coherent.

HTML Hypertext Markup Language is the set of markup symbols inserted in a file intended for display on a World Wide Web browser. The markup instructs the Web browser how to display a Web page.

IIOP Internet Inter-ORB Protocol— a standard that ensures interoperability for objects in a multi-vendor ORB environment.

Integrity In a client/server environment, integrity means that the server code and server data are centrally maintained and therefore secure and reliable.

Information Machine readable content or data that is in the correct format to be processed by an application or system.

Invasive Integration An implementation approach that requires changes or additions to existing applications. Opposite of non-invasive integration.

IOTP Internet Open Trading Protocol. This is a set of standards developed by the IETF to define an interoperable framework for Internet commerce.

JVM: Java virtual machine.

J2ME: Java 2 platform Micro Edition.

Load Balancing Automatic balancing of requests among replicated servers to ensure that no server is overloaded.

LU6.2 Logical Unit 6.2— IBM’s device-independent process-to-process protocol provides the facilities for peer-to-peer communications between two programs and also supports asynchronous networking.

Message Broker A key component of EAI, a message broker is an intelligent intermediary that directs the flow of messages between applications, which become sources and consumers of information. Message brokers provide a very flexible communications backbone and provide such services as data transformation, message routing and message warehousing.

MOM Message-Oriented Middleware is a set of products that connects applications running on different systems by sending and receiving application data as messages. Examples are RPC, CPI-C and message queuing.

Message Queuing A form of communication between programs. Application data is combined with a header (information about the data) to form a message. Messages are stored in queues, which can be buffered or persistent (see Buffered Queue and Persistent Queue). It is an asynchronous communications style and provides a loosely coupled exchange across multiple operating systems.

Message Routing A super-application process where messages are routed to applications based on business rules. A particular message may be directed based on its subject or actual content.

Message Warehousing A central repository for temporarily storing messages for analysis or transmission.

Middleware Software that facilitates the communication between two applications. It provides an API through which applications invoke services and it controls the transmission of the data exchange over the network. There are three basic types: communications middleware, database middleware and systems middleware.

Mobile Distributed Objects: (MDOs). A distributed object that can migrate from one processor or process to another (aka Jini).

Non-Blocking Communications An asynchronous messaging process whereby the requestor of a service does not have to wait until a response is received from another application.

Non-invasive: The quality of not requiring code modifications, not imposing unnatural communication protocols or protocol semantics, and not degrading performance or quality.

Non-Invasive Integration An implementation approach that does not require changes or additions to existing applications.

Open Applications Group (OAG) An industry consortium formed to promote the easy and cost-effective integration of key business application software components.

Object Middleware Allows clients to invoke methods or objects that reside on a remote server. This middleware revolves around OMG's CORBA and Microsoft's DCOM.

ODBC Open Database Connectivity— a Windows standard API for SQL communication.

OMG Object Management Group— a consortium of object vendors and the founders of the CORBA standard.

OpenDoc A set of shared class libraries with platform-independent interfaces.

ORB Object Request Broker— software that allows objects to dynamically discover each other and interact across machines, operating systems and networks.

PDA: Personal Digital Assistant.

Persistent Queue A message queue that resides on a permanent device, such as a disk, and can be recovered in case of system failure.

Publish/Subscribe Pub/Sub is a style of inter-application communications. Publishers are able to broadcast data to a community of information users or subscribers, which have issued the type of information they wish to receive (normally defining topics or subjects of interest). An application or user can be both a publisher and subscriber.

Quasi-Synchronous Communication: An interaction between two autonomous entities in which the initiator does not suspend execution, but does set forth a mechanism to catch the recipient's response.

RDA Remote Data Access, usually to an RDBMS via SQL.

RDBMS Relational Database Management System.

Router A special-purpose computer or software package that handles the connection of two or more networks.

Routers check the destination address of the packets and decide the route to send them.

RPC Remote Procedure Call— a form of application-to-application communication that hides the intricacies of the network by using an ordinary procedure call mechanism. It is a tightly coupled synchronous process.

Scalability The ability of an information system to provide high performance as greater demands is placed upon it, through the addition of extra computing power.

Server A computer or software package that provides specific capabilities to client software running on other computers.

SNA System Network Architecture— a network architecture from IBM.

Sockets A portable standard for network application providers on TCP/IP networks.

SQL Structured Query Language.

Stored Procedure A program that creates a named collection of SQL or other procedural statements and logic that is compiled, verified and stored in a server database.

STP Straight Through Processing. This occurs when a transaction, once entered into a system, passes through its entire life cycle without any manual intervention. STP is an example of a Zero Latency Process, but one specific to the finance industry which has many proprietary networks and messaging formats.

Supply Chain Management Supply Chain Management (SCM) is that set of skills and disciplines, including those of IT, which shepherd a product from its original design to its ultimate delivery to the buyer.

Synchronous Communications A form of communication that requires applications to run concurrently. A process issues a call and until it receives a response. (Or) An interaction between two autonomous entities in which the initiator suspends all subsequent execution after the

communication event until it receives some response (e.g., return value).

Systems Middleware Software that provides value-add services as well as inter-program communications. An example is transaction processing monitors which are required to control local resources and also cooperate with other resource managers to access non-local resources.

TCP/IP Transmission Control Protocol/Internet Protocol—the network protocol for the Internet that runs on virtually every operating system. IP is the network layer and TCP is the transport layer.

Trigger A stored procedure that is automatically invoked on the basis of data-related events.

Two-Phase Commit A mechanism to synchronize updates on different machines or platforms so that they all fail or all succeed together. The decision to commit is centralized, but each participant has the right to veto. This is a key process in real time transaction-based environments.

VPN Virtual Private Network. A solution by which an enterprise may link to its customers and business partners via secure Internet connections. The company thus has a network which is "just like" a private network, but isn't. (That is, it is "virtual.") This gives the company the advantages of a private network at the much lower cost of a public one.

Workflow Software used to automatically route events or work-items from one user or program to another. Workflow is synonymous with process flow, although traditionally has been used in the context of person-to-person information flows.

XML Like HTML, eXtensible Markup Language is a subset of Standard Generalized Markup Language, a standard for defining descriptions of structure and content in documents. However, where HTML is concerned with the presentation of information on a web page (without context or dynamic behavior), XML provides context and gives meaning to data.

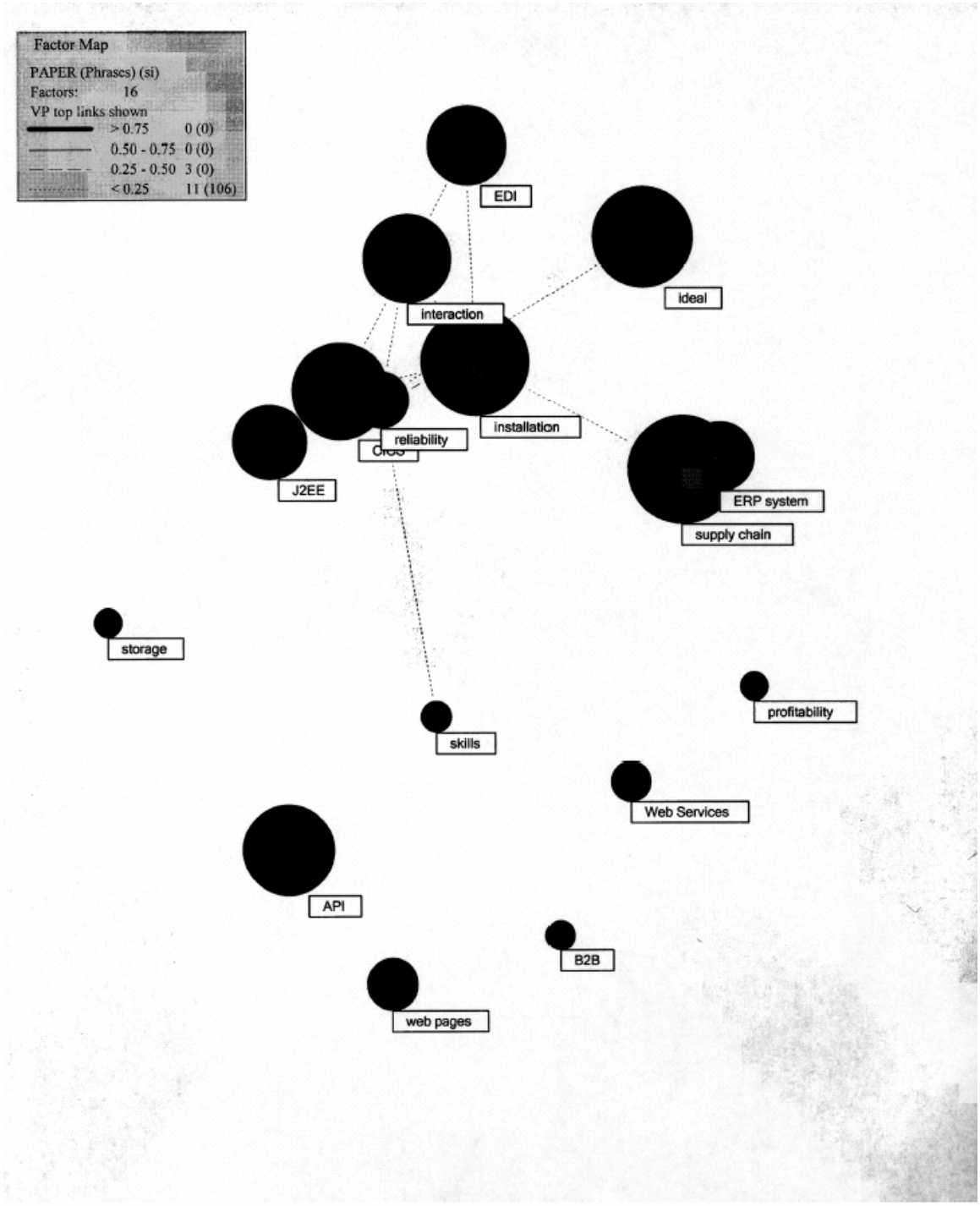
X/Open An independent open systems organization. Its strategy is to combine various standards into a comprehensive integrated systems environment called Common Applications Environment, which contains an evolving portfolio of practical APIs.

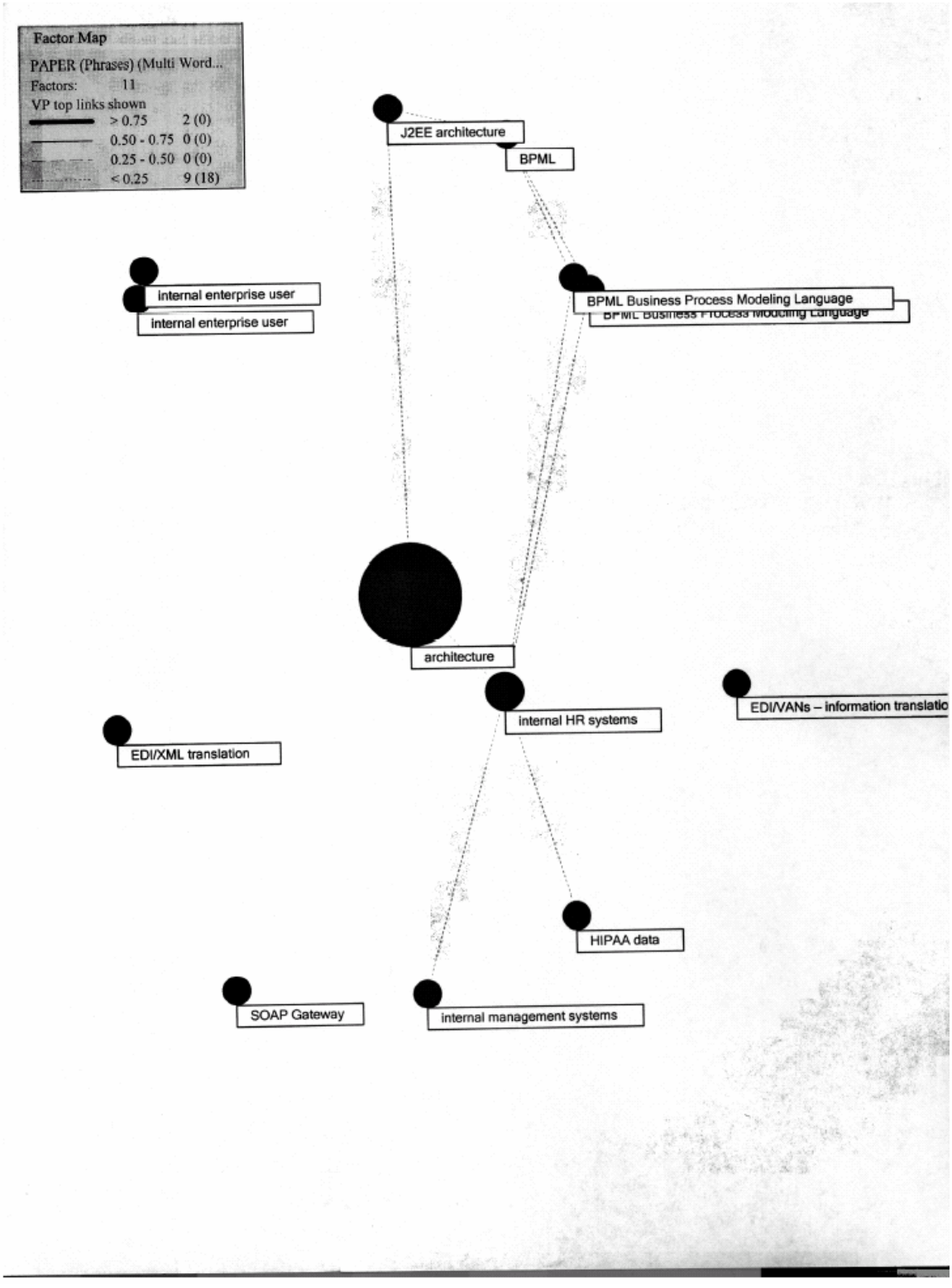
Zero Latency No delay between an event and its response.

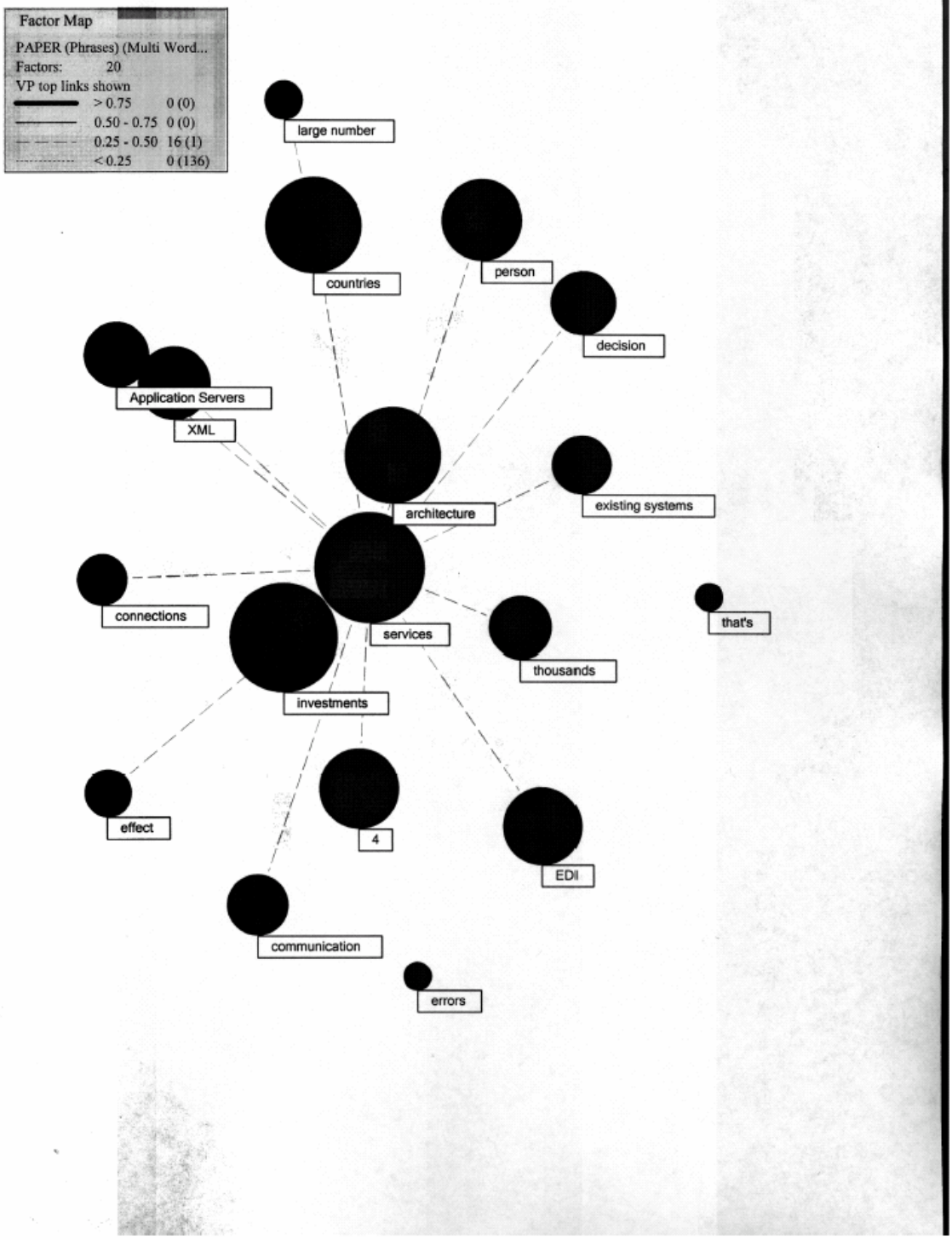
Zero Latency Enterprise An enterprise in which all parts of the organization can respond to events as they occur elsewhere in the organization, using an integrated IT infrastructure that can immediately exchange information across technical and organizational boundaries.

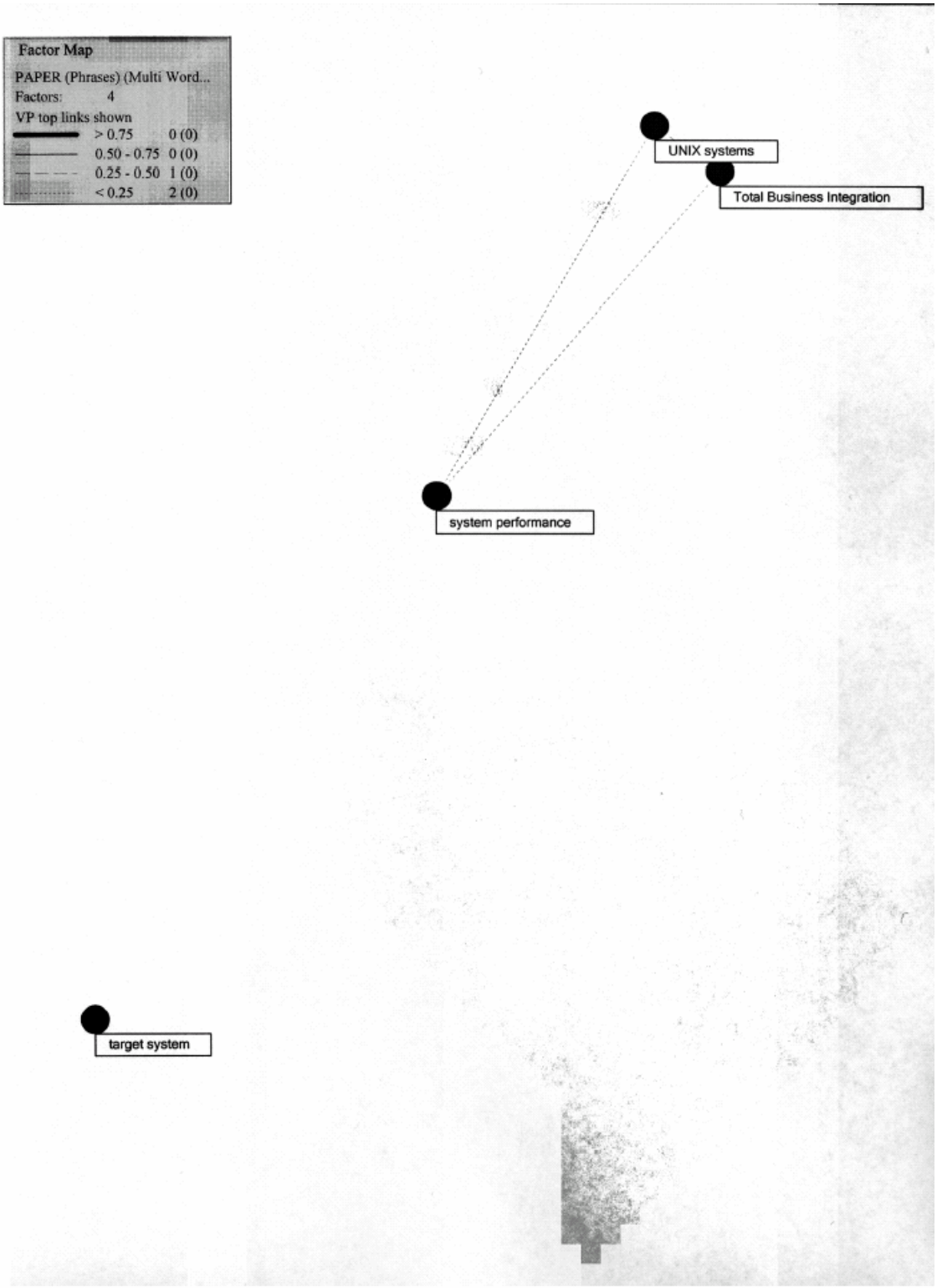
Zero Latency Process An automated process with no time delays (i.e. no manual re-entry of data) at the interfaces of different information systems. STP is an example.

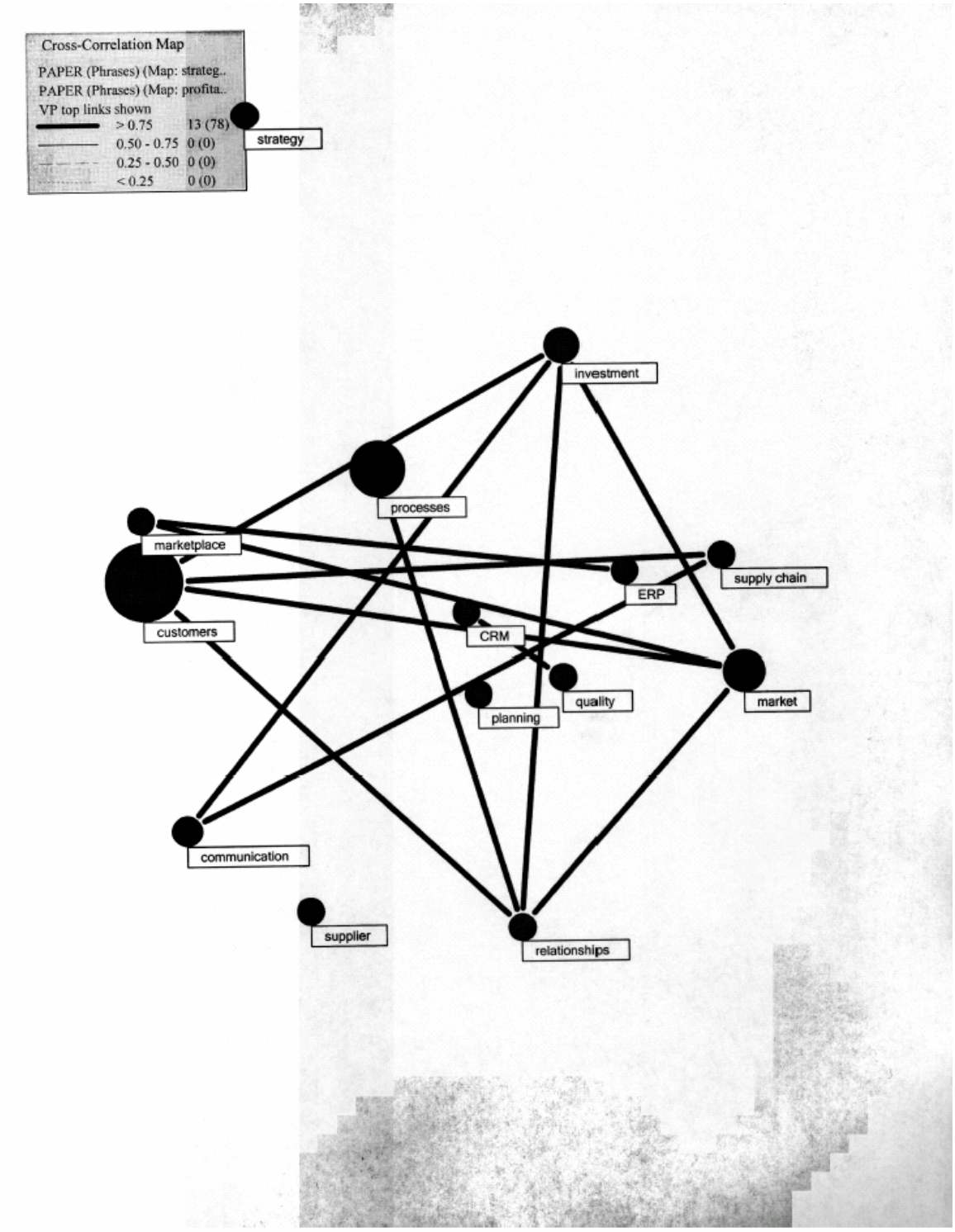
APPENDIX B: VANTAGE POINT OUTPUTS

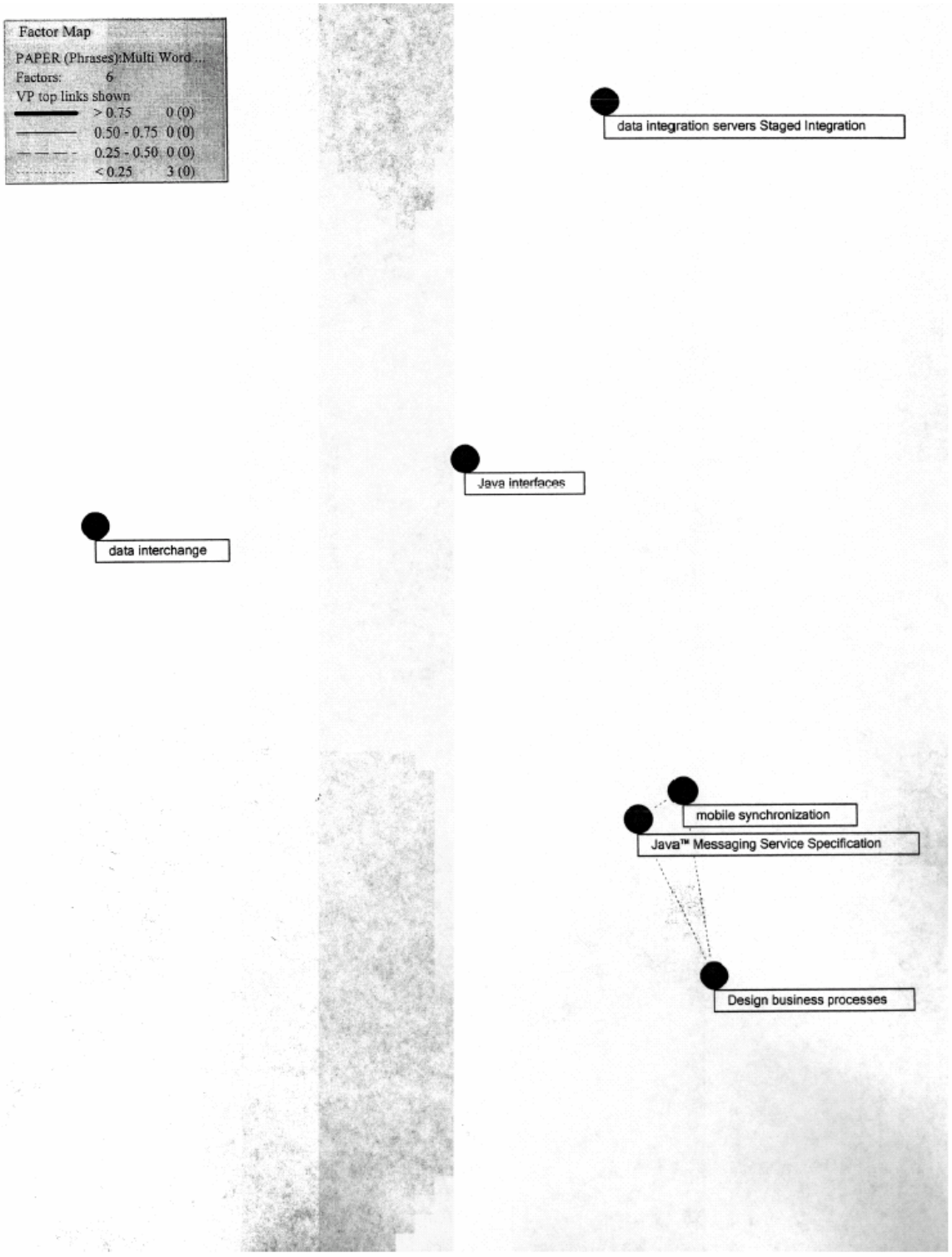

Factor Map
PAPER (Phrases) (si)
Factors: 16
VP top links shown
> 0.75 0 (0)
0.50 - 0.75 0 (0)
0.25 - 0.50 3 (0)
< 0.25 11 (106)
EDI
ideal
interaction
installation
reliability
J2EE
ERP system
supply chain
storage
profitability
skills
Web Services
API
B2B
web pages

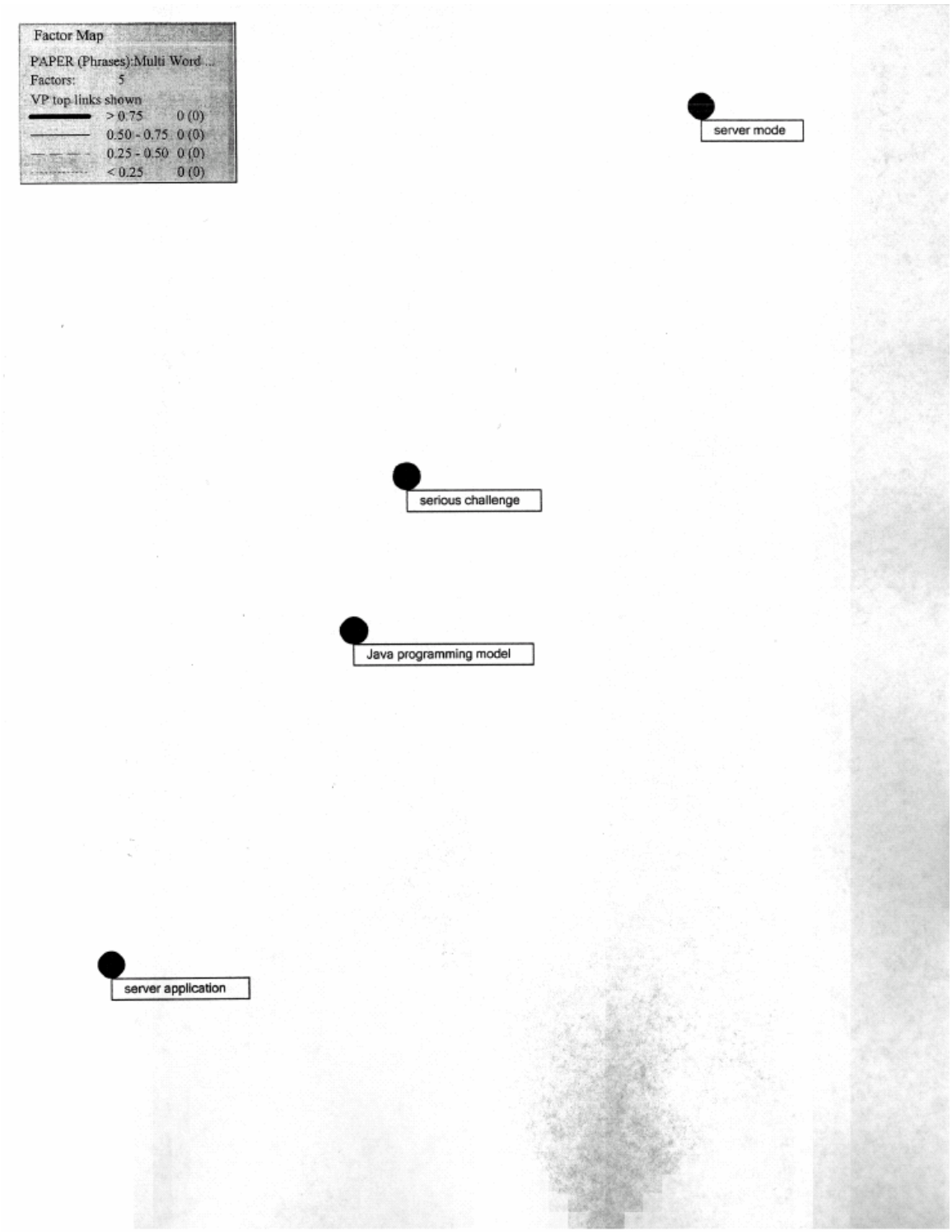

Factor Map
PAPER (Phrases) (Multi Word...
Factors: 11
VP top links shown
> 0.75 2 (0)
0.50 - 0.75 0 (0)
0.25 - 0.50 0 (0)
< 0.25 9 (18)
J2EE architecture
BPML
internal enterprise user
internal enterprise user
BPML Business Process Modeling Language
architecture
internal HR systems
EDI/VANs – information translatio
EDI/XML translation
HIPAA data
SOAP Gateway
internal management systems


Factor Map
PAPER (Phrases) (Multi Word...
Factors: 20
VP top links shown
> 0.75 0 (0)
0.50 - 0.75 0 (0)
0.25 - 0.50 16 (1)
< 0.25 0 (136)
large number
countries
person
decision
Application Servers
XML
architecture
existing systems
connections
services
thousands
that's
investments
effect
4
EDI
communication
errors


Factor Map
PAPER (Phrases) (Multi Word...
Factors: 4
VP top links shown
> 0.75 0 (0)
0.50 - 0.75 0 (0)
0.25 - 0.50 1 (0)
< 0.25 2 (0)
UNIX systems
Total Business Integration
system performance
target system


Cross-Correlation Map
PAPER (Phrases) (Map: strateg..
PAPER (Phrases) (Map: profita..
VP top links shown
> 0.75 13 (78)
0.50 - 0.75 0 (0)
0.25 - 0.50 0 (0)
< 0.25 0 (0)
strategy
investment
processes
marketplace
supply chain
ERP
customers
CRM
quality
market
planning
communication
supplier
relationships


Factor Map
PAPER (Phrases):Multi Word ...
Factors: 6
VP top links shown
> 0.75 0 (0)
0.50 - 0.75 0 (0)
0.25 - 0.50 0 (0)
< 0.25 3 (0)
data integration servers Staged Integration
Java interfaces
data interchange
mobile synchronization
Java™ Messaging Service Specification
Design business processes


Factor Map
PAPER (Phrases):Multi Word ...
Factors: 5
VP top links shown
> 0.75 0 (0)
0.50 - 0.75 0 (0)
0.25 - 0.50 0 (0)
< 0.25 0 (0)
server mode
serious challenge
Java programming model
server application

VITA

Balaji Ravindran was born and brought up in Madras, India. He has an undergraduate degree in *Civil Engineering* from Bangalore University. His interests in software eventually led him to pursue a career as *Software Developer and Tester* for 3 years prior to Master's degree in *Engineering Management* from University of New Orleans. His career interests include *Enterprise Integration* consulting, and writing articles in the same field for technical magazines.