KENT STATE
UNIVERSITY

Department of Management and Information Systems

Working Paper Series

A Systemic Ontology of Service-Driven IT Infrastructure Design

Pratim Datta

Working Paper **MIS2007:14**

Department of Management and Information Systems, College of Business, Kent State University, Kent, OH 44242, USA

A SYSTEMIC ONTOLOGY OF SERVICE-DRIVEN IT INFRASTRUCTURE DESIGN

Pratim Datta (corresponding author)
Kent State University
Department of Management and Information Systems
A408 College of Business
Kent, OH 44242
330-672-1229 (voice)
509-335-4275 (fax)
pdatta@wsu.edu

ABSTRACT

The purpose of this paper is to introduce a new conceptualization of information technology (IT) infrastructure based on the principles of general systems theory. The IT infrastructure is the overarching IT artifact in organizations: complex and dynamic. Yet, there is no evidence of conceptual models that can recognize and address the reconfiguration and evolution of an IT infrastructure in an organization. By shifting our theoretical lens to the modular systems perspective, we use key attributes of modularity, integration, (re)configuration, and feedback to extend existing notions and forwarding a dynamic model of IT infrastructure. Instead of assuming IT infrastructure as a static installed base of technologies, we try to map IT infrastructure as being-service driven, constantly reorganizing and balancing its technology and human assets to adapt to changing service demands. To illustrate the value of this approach, we use an open systems theory lens to classify corporate IT infrastructure into segmented, shared, and seamless infrastructure assets, configurable on-demand. By rethinking IT infrastructure as a portfolio of configurable subsystems, the model explains how corporate IT infrastructures constantly reorganize and adapt themselves between differentiation and integration.

Keywords: Modular Design, IT Infrastructure, Open Systems Theory, Integration

> *"Seventy-five percent of all IT dollars go to infrastructure. Isn't it time you learned what it is?"*
>
> *(IBM advertisement cf. Applegate et al. 2006).*

1. INTRODUCTION

An organization’s information technology (IT) infrastructure is constantly evolving. The evolution is rapid and recurrent: legacy systems are bridged, networks are revamped, applications are integrated, personnel are cross-functionally deployed, and new services are defined and assimilated. While notable advances have been made towards conceptualizing IT infrastructure (e.g., Broadbent et al., 1996; Weill and Broadbent, 1998; Weill and Vitale, 2002), the purview of their arguments has been captive to treating IT infrastructure as a stable and monolithic artifact consisting of discrete, insulated resources. IT infrastructure has long been understood in terms of insulated resources such as processors, servers, storage, routers, operating systems, applications and services. However, the need for IT infrastructure agility and flexibility has impelled organizations towards integrating their infrastructure to match a variety of overlying reengineered business processes. As a result, IT infrastructure over time has become “complex ensembles of heterogeneous artifacts, which are increasingly connected with and dependent upon one another” (Hanseth and Lyytinen, 2005: 1). Unfortunately, few IT infrastructure design models exist that encapsulate the evolution of IT infrastructure from insulation to integration, particularly when driven by growing concerns of agility, scalability, and flexibility.

Despite advances made by past research in explaining IT infrastructure, existing research has been captive to traditional silo-based functional IT artifacts while the overarching dynamics of the IT infrastructure has generally escaped scrutiny (Broadbent et al., 1996, 1999). Shifting and evolving IT infrastructure increases in complexity. Organizations run standalone legacy and integrated systems in parallel. Often, thanks to application programming interfaces and interface definition languages, technology assets are being assigned for providing different services within the organization. For example, while some mainframes are being used for standalone computing, others are being reassigned as content servers; while some legacy systems are being maintained for their functional efficiencies, other legacy systems are being rehauled through device-level and operating and application levels updates. Simply put, companies organize IT infrastructures in novel ways: catering to diverse demands to remain competitive. Yet, archetypes actually addressing either the novelty or diversity of IT infrastructure are rare. If

IT infrastructure is perplexing, it is because we do not have a good model to apply to its analysis (Fomin, 2003; Ciborra et al., 2001). In short, current conceptualizations suffer from certain limitations. Most existing models treat corporate IT infrastructure as insular, static, overly simplistic, supply-centric, and lacking adaptation or fit. Such limitations lie at the root of our inability to understand and portray the IT infrastructure artifact.

Contending that the evolving dynamics of IT infrastructure requires a fresh lens for scrutiny and clarification, this paper is an attempt to capture and model the design of IT infrastructure in organizations. While prior conceptual framings have been effective in instituting a groundwork for analysis, they do not account for the dynamic relationships and interdependencies between infrastructure elements. Nor do they offer clarification in regards to the adaptation and evolution of a firm's IT infrastructure over time. Without a systematic conceptualization of IT infrastructure assets, IT infrastructure becomes a potentially confounding mix of resources spread across the firm (Weill and Broadbent, 1998), resembling "a somewhat of a messy collage, as a result of deals, improvisations, and layers of sedimentation" (Ciborra et al., 2001: 3). To address these limitations, we propose a new systems-based reconceptualization of IT infrastructure capable of capturing the dynamics of an organization's IT infrastructure configurations.

The paper is organized as follows: We begin with a review of prior conceptual framings of IT infrastructure. Next, we turn to general systems theory, beginning with a description of basic concepts central to the reframing of IT infrastructure. Systems theory is then used in combination with insights from prior research to forward our integrated model, exemplifying how systems theory offers clarification and conceptualization benefits. Lastly, the framework is utilized to derive a model of firm IT infrastructure modularity, demonstrating the value of an open systems view of infrastructure in developing new theoretical insights. The paper concludes with a discussion of the implications of the re-conceptualization for research and practice.

RETHINKING IT INFRASTRUCTURE

The reconceptualization of IT infrastructure begins with a definition aimed at reducing the aforementioned limitations of prior conceptualizations. IT infrastructure is defined as *a service-driven heterogeneous portfolio of modular and configurable assets consisting of technologies, applications, people and IT-enabled routines (processes)*. This definition emphasizes three key elements: (a) IT infrastructure is service-driven; (b) IT infrastructure is a portfolio of modular and configurable assets; and (c) IT infrastructure encapsulates technologies and people assets.

The definition reframes IT infrastructure as a key enabler instead of a base, rigid substrate of technologies. As Hanseth and Lyytinen (2005) argue, IT infrastructure is capability oriented. IT infrastructure must be capable to offer requisite services for objective benefits. Departing from an old supply-push adage that services follow IT infrastructure, we find it more relevant to consider a demand-pull effect where services drive IT infrastructure. The supply-push notion seems appropriate for the mainframe-era, where corporate data center services were deployed from, and thus captive to, mainframe assets (Applegate et al, 2006). However, this often led to the development of core rigidities, where the ossified captivity of services to underlying IT infrastructure assets makes the IT infrastructure inflexible and intractable to shifts in the competitive landscape. However, as IT grows to be more pervasive and ubiquitous, service-delivery becomes key - forcing reconfigurations of IT infrastructure to maintain pace with changing business models. For instance, customer relationship management is a different service delivery model from service delivery assurance and would require a different configuration (and, perhaps, reconfiguration) of IT infrastructure assets. However, because a firm involves itself in multiple service activities, it ends up with a bundle of ad-hoc IT infrastructure asset configurations. Modularity gives IT architects and corporate managers the freedom to consider the pieces (assets) that make up IT infrastructure and their configuration (fit) to the overall business services-delivery strategy. The ability to rethink IT infrastructure as a modular bundle of assets (Weill and Broadbent, 1998), i.e., technology, people and processes, provides the freedom to add, combine, and remove IT infrastructure assets as independent, manageable pieces that can be matched and melded for IT infrastructure scalability and flexibility. According to Applegate et al. (2006: 473), “as service delivery models proliferate and improve, the variety of IT asset configurations will increase.” Reframing IT infrastructure as a service-driven modular and configurable portfolio of assets allows firms to understand how to best deploy IT infrastructure assets for maximum return.

Employing such a perspective offers tremendous benefits for both IT and non-IT executives. First, because information systems (e.g., billing systems, purchasing systems) are a function of their services, a service driven IT infrastructure asset framework would explain how organizations configure their IT infrastructure assets to provide these services. Second, viewing IT infrastructure as a modular and configurable portfolio of assets simplifies the executives' ability to assess, secure and control IT infrastructure assets for objective assurances (e.g., following CobIT, ITIL). Third, given the proliferation of shared IT services, understanding how services drive IT infrastructure asset configurations can be useful for streamlining asset management and total cost of ownership (TCO). Fourth, a configurable asset design simplifies the complexity of IT infrastructure to better understand reusability, convergence, and adaptability. Fifth, a configurable asset view of IT infrastructure is no longer captive to technology but encompasses people and processes - reconfiguring technology, retraining people and reengineering processes. Finally, as firms go through mergers and acquisitions or outsource service, the proposed IT infrastructure design can easily accommodate various degrees of infrastructure asset consolidations and reductions owing to mergers, acquisitions, and outsourcing. In our knowledge, such an integrative approach has been, hitherto, absent in mainstream IT literature.

II. FRAMING IT INFRASTRUCTURE – PRIOR WORK AND PRESENT CHALLENGES

Although various views of IT infrastructure have been much debated in the past, currently there is a high degree of agreement within the IS research community regarding the general scope. A seminal attempt was the Zachman Framework for Information Systems Architecture (Zachman, 1987) that provided logical and physical specifications to define and control component-level interfaces and integration. While Zachman's framework provides a consolidated view of information systems at different levels of scope (e.g., data, process, network) and detail (e.g., designer view, developer view, owner view), it has often been criticized as being too complex and confusing. In short, although Zachman's framework adds value to understand the overall layers of technology organization and operations, the framework does not account for the diversity of IT infrastructure at different levels of integration or convergence nor does it provide a dynamic view of infrastructure.

Delineation of the scope of IT infrastructure is further captured by a synthesis by Byrd and Turner (2000), who, building on the layered model by Broadbent et al. (1996), defined IT infrastructure as shared resources consisting of a technical physical base (e.g., hardware, software, networks, data) and human components (e.g., skills, expertise, knowledge, commitments) combined to create IT services. A common approach in these studies has been the delineation of IT infrastructure in terms of: 1) *technical* IT infrastructure (e.g., networks, applications, platforms, data), and 2) *human* IT infrastructure (e.g., technology management, business functional skills, interpersonal skills, technical specialty skills) (Duncan, 1995; Lee et al., 1995). The layered model was seminal in forwarding the presence of an application core and a human element on top of an 'installed base' of technology.

Although the layered model has become the predominant framework for IT infrastructure in IS research today, it is problematic in two respects. The first relates to the specifications of two of the layers that are used to define the technical IT infrastructure: 1) IT components, and 2) shared and standard IT applications. IT components are construed as a conglomeration of technologies, including databases, routers, operating systems, monitors, etc. The lack of classification of these components presents a severe limitation when trying to evaluate an IT infrastructure, identify specific areas of investment, and so on. Putting monitors and hard drives in the same placeholder as that of an enterprise application system seems unjustified and misleading. The second issue relates to shared and standard applications. The simplifying assumption reflected in the model is that the types of organizational applications of IT cited are stable and standardized although, in reality, for most firms they are not (Lewis and Byrd, 2003). Many applications are still relegated to functional departments or processes, leading to a multiplicity of standards and limited sharing (Kayworth et al., 1997). Moreover, application of technologies is rarely standardized, as users tend to reorient themselves differently, leading to distinct perceptions and applications of the same technology (Orlikowski, 1993). For example, there are firms where one function uses Microsoft Excel to maintain accounting ledgers, while another uses it to control inventory, while the third may use it to develop VB applications. Using examples of large-scale enterprise applications such as ERP systems, Weill and Vitale (2002) note the current trend to standardize. However, they fail to capture the fact that IT infrastructure is still a portfolio of technologies at various degrees of convergence within most organizations.

Going beyond these layered, descriptive models, Ciborra et al. (2001) were elemental in studying the evolutionary dynamics of corporate IT infrastructure by "moving infrastructure from a thrown-together institutional backbone to a value-generating, integrated set of technologies, applications, and processes" (p. 3). Central to Ciborra et al.'s argument was that IT infrastructures tend to 'unintendedly' drift from their planned purposes. Using Giddens' concept of modernity, Ciborra et al. offer multiple critical views (e.g., using the actor network theory) to portray how IT infrastructure evolves independent of assertions and the credo of bureaucratic and management control. By marrying theories with a portfolio of case studies, the authors provide essential insights into how technology is more than just a passive tool. Instead, as the authors find, IT infrastructure is dynamic and drifting: shaping organizations and shaped by control and interpretations. While Ciborra et al.'s (2001) arguments are extremely useful in understanding the shaping of corporate IT infrastructure and begin to address some of the concerns raised earlier in the paper, there is not systematic clarification of corporate IT infrastructure itself.

The use of socio-technical ethnographic perspectives to understand IT infrastructure (e.g., Edwards, 2003; Star, 1999; Monteiro, 2001) further advances our understanding of IT infrastructure evolution. The socio-technical view considers how a network of multiple constituencies (e.g., manufacturers, users, service providers) influences IT infrastructure. Star, for example, views IT infrastructure as relational, shaped by organized practices - a crystallization of institutional relationships embodying power and control. This perspective, rich in insights, shifts the locus of investigation from the technological to the socio-political collective (Edwards, 2003; Bowker and Star, 1999) - making IT infrastructure relational and ecological (Star, 1999). The use of actor network theory (ANT) builds on the socio-technical ethnography to further our understanding of IT infrastructure (Monteiro, 2001). ANT builds its arguments on the premise that an actor is a part of a web of entities, technical or non-technical, tied together as an interdependent and evolving network. Different nodes influence and shape the actor. The actor translates these influences from its participating network to inscribe and embody them into certain artifacts aimed to trigger certain outcomes. The inscription evolves though trials, errors and the translation of the overarching influences. The inscription is successful only when it aligns with the translations of other actors in the network. For example, Monteiro (2001) uses ANT to how infrastructure is actively recursive rather than a passive artifact - how, by acting as an installed base and platform, IT infrastructure pulls and is pulled by socio-political and socio-technical networks. By explaining IT infrastructure evolution as an outcome shaped by a web of participating relationships, ANT enhances the socio-technical perspective forwarded by Star (1999). However, perceiving IT infrastructure as an evolving web of relationships heightens the complexity of systematically translating IT infrastructure for management.

Although previous advances have tremendously contributed towards explaining IT infrastructure, a common limitation has been scant attention towards a systematic representation of a scalable, flexible, and service-driven infrastructure. Such a perspective offers limited visualization, failing to capture the evolving realities surrounding corporate infrastructure. Given the need for a systematic representation of the dynamics of IT infrastructure, it is important to rely on theories that can help us understand the evolving and adaptive nature of service-driven IT infrastructure at various levels of configuration and integration. This is the topic we turn to in the next section.

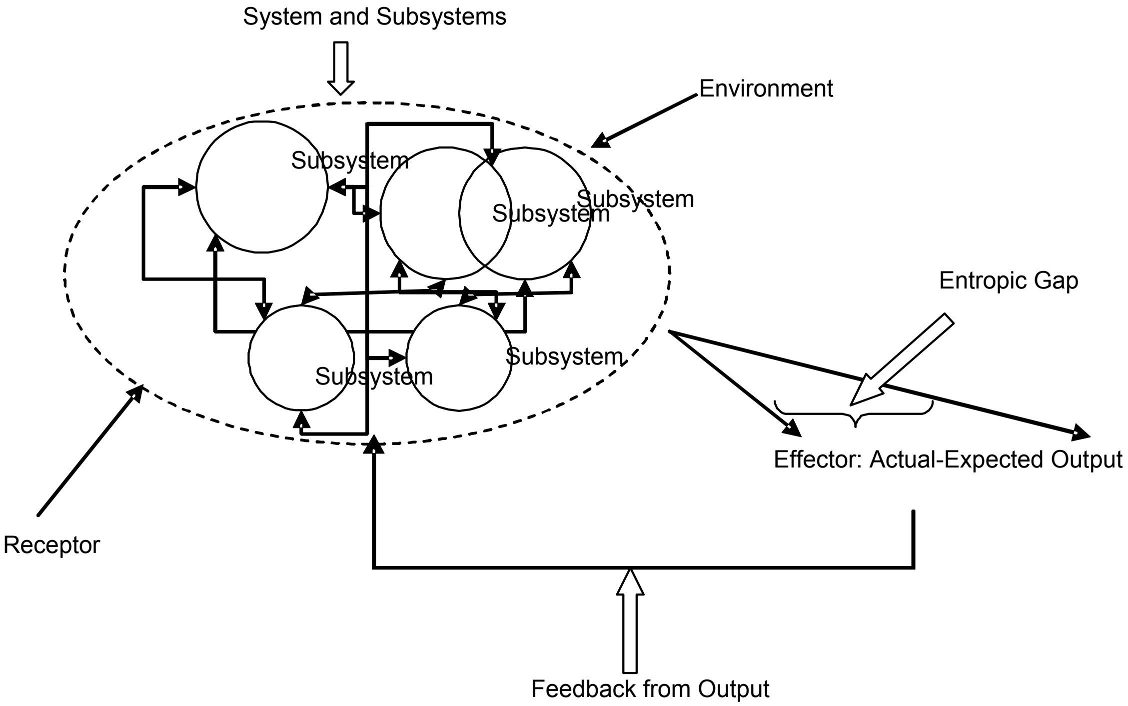
III. REFRAMING IT INFRASTRUCTURE – A SYSTEMIC ONTOLOGY

"The only meaningful way to study organization is to study it as a system," remarks Scott (1961: 22). In *The Sciences of the Artificial,* Herbert Simon (1981) argues for the need to characterize artificial, or man-made, artifacts (e.g., IT infrastructure) as hierarchically nested "open" organizational systems that serve a function, help achieve a goal and adapt in alignment with the organization. Although a concept originally developed for studying organisms, systems theory has been used to study any complex adaptive system. The systemic approach treats systems as generally open, i.e., it is adaptive to the demands of the environment, and in turn, often shapes the environment. A system builds on "differentiation and integration" (Lawrence and Lorsch, 1967). On one hand, a system is reducible into hierarchies of nested subsystems to match the examiner's expected level of abstraction. On the other hand, multiple modular subsystems can be integrated by virtue of their levels of interdependencies and interactions among themselves or with their environments. This simultaneous presence of cohesion and differentiation creates a Janus (the Roman two-headed god) effect where a system can has multiple subsystems, but bound by synergies. Nested hierarchies, interactions, openness, together contribute toward the systems' and subsystems' goal-seeking behavior.

The word 'system' is derived from the Greek word *'synistanai,'* which means to bring together or combine. First proposed by von Bertalanffy (1968) in the 1940's, general systems theory "is the investigation of organized wholes" (p. 32). Ashby (1956), in his seminal article on cybernetics, considers systems as an observer's preferred description of a set of interrelated elements connected by an organized stream of information, maintaining "independence within a whole" (p. 55). Such a system can exist at multiple levels of abstraction and complexity. Boulding's (1956) influential paper is perhaps the single most important introduction of systems theory in the organizational context, placing systems theory as a balance between the overly abstract and the overly specific: "In recent years increasing need has been felt for a body of systematic theoretical constructs which will discuss the general relationships of the empirical world. This is the quest of General Systems Theory…between the specific that has no meaning and the general that has no content there must be, for each purpose and at each level of abstraction, an optimum degree of generality." (p. 197) Systems theory has thus "come into use to describe a level of theoretical model-building which lies somewhere between the highly generalized constructions of pure mathematics and the specific theories of the specialized disciplines" (Boulding, 1956: 197).

The open systems approach views organizations as purposive systems, emphasizing differentiation, integration, and interaction (Scott, 1998). As shown in Figure 1, an organization consists of a systemic process of input-throughput (process/transformation)-output, with receptors (input), processors (for reconfiguration and transformation), and effectors (output). In contrast with cybernetics where system processes are treated as a black box, systems theory examines all essential processes. A system comprises a set of constituent systems or subsystems. Levels of subsystems can be considered infinite, open to the level of granularity that an investigator seeks. The ability to decompose a system into subsystems or modules is the essence of modularity, as subsystems can change their configurations over time. While subsystems are autonomous in form, they are cohesive in function, i.e., these relatively self-contained subsystems interact with one another to serve a unified objective (Emery and Trist, 1965). The boundary defines the scope of the system (e.g., IT infrastructure) or the subsystem. Beyond the boundary is the system environment consisting of opportunities, demands and constraints. A system receives its input from the environment and sends its output to the environment. The difference between the expected and the actual output is referred to as the entropic gap, fed back to the system. A system attempts to reduce entropy by matching expected to actual output. Commonly, systems try to resolve the entropic gap through feedback mechanisms. Systems use feedback mechanisms to reconfigure subsystems. The effector maps the systems output and the information is fed back to the controller. The feedback serves as a control mechanism for maintaining homeostasis, a phenomenon drawn from biology where negative feedback is used to control undesirable variations and positive feedback is used to induce desirable variations in performance-accommodating continuous change and adaptation.

Figure 1. The Open Systems Perspective



The objective of our use of systems theory as a theoretical lens is to richly depict the evolving dynamics of a scalable, flexible corporate IT infrastructure. Applying traits used to define an open system adds significant value in understanding a service-driven IT infrastructure. Concepts from the open systems theory central to understanding IT infrastructure are *subsystems (modularity), receptors, effectors, controllers (for differentiation and integration), environment (constraints), entropic gap, and homeostasis (adaptation) and feedback.*

A *subsystem* is a lower level of hierarchy assigned to any system, often a function of the observers' choice of abstraction (e.g., metasystem, system, subsystem, component). Subsystems allow reducing a system into more granular modules for purposes of understanding and analysis (Skyytner, 2001; van Gigch et al., 1981). Subsystems denote modularity - a degree to subsystems can be separated and recombined based on rules set forth by the system controller (Schilling, 2000). Modularity therefore “exponentially increases the number of possible configurations achievable from a given set of inputs” (Schilling, 2000: 4). In relation to IT infrastructure, subsystems are modular assets (e.g., technologies, people, processes) that constitute the IT infrastructure portfolio. Modularity can thus underpin the coupling and recombination of IT infrastructure subsystem assets to create a mix for optimal service.

A *receptor* allows a system or subsystems to detect and receive input from the environment (van Gigch, 1981). A receptor binds the system or a subsystem to a context, receiving and registering ad-hoc demands for processing (Skyytner, 2001). In the context of IT infrastructure, receptors receive service-requests from users and other stakeholders, validate the service request (e.g., user authentication, privileges), and pass it to be processed by the IT infrastructure. In IT infrastructure systems, the receptor would most likely be a service gateway.

An *effector* outputs information processed by the controller. Effectors are goal-oriented, tracking actual versus expected output, and triggering feedback to the controller as required (van Gigch, 1981). In IT infrastructures, effectors compare service outputs from the IT infrastructure to service expectations. An effector continuously reports variations to reconfigure and reorganize subsystems for alignment and fit. The aim of the effector is to reduce the level of disorder or *entropy* within the IT infrastructure system by comparing variations from the expected IT infrastructure service outcomes and providing requisite feedback to the system or subsystems for reconfiguration.

A *controller* is the transformation and processing core of the system (Skyytner, 2001). Controllers regulate internal systems (subsystems) based on ad-hoc inputs. Controllers dynamically configure and reconfigure subsystems based on certain service demands and organizational goal-orientation. In IT infrastructure, controllers organize and regulate IT infrastructure subsystems to cater to service requests from the receptor. According to the goals and constraints of specific services, the controller can organize and regulate IT infrastructure systems towards increasing (differentiation) or decreasing modularity (integration).

An *environment* is the surrounding *milieu* within which an open system operates.

Unlike a closed system that is defined by a fixed set of inputs and outputs, an open system is always dependent on an environment. The system and its subsystems use receptors and effectors to exchange information with their environments. In the context of IT infrastructure, the environment can be the industry, market, or the socio-technical network of stakeholders (e.g., Ciborra, 2001). Corporate IT infrastructures can change and be changed by the environment. For example, the environment may demand certain services from the IT infrastructure (e.g., policy-driven IS assurance services); at the same time, IT infrastructure can also change the environment with service innovations (e.g., P2P services).

Homeostasis is the attempt of a system to maintain equilibrium through control and regulation. Purposive control and regulation allows the system to constantly reconfigure and reorganize its subsystems and their interactions to maintaining a degree of equilibrium. Central to maintaining homeostasis is f*eedback* - a mechanism that allows a system to compensate for disturbances in the environment by reconfiguring subsystem assets. The idea of homeostasis through feedback is what makes a system adaptive (Skyytner, 2001; van Gigch et al., 1981). IT infrastructure, as a system, maintains homeostasis by adaptively reorganizing subsystems to meet specific service demands. Specific services demand the employment of certain combinations of subsystems - the controller adaptively organizes these subsystems in various degrees of modularity to match service demands to reduce entropy. Information on entropic gaps (large levels of variation) in expected versus service outputs is fed back by effectors to the controller that consequently regulates the subsystems to maintain equilibrium. A negative feedback loop is a negative multiplier while a positive feedback is a positive multiplier. Based on service outputs from a previously modular system, a negative feedback loop will reduce modularity (increase integration) while a positive feedback loop will increase modularity.

To summarize, by making the environment a part of the overarching system, open systems theory captures the dynamics from the mutual interactions between the system and its immediate environment, therefore addressing, in many ways, the central themes forwarded by ethnographic and actor-network perspectives used to build IT infrastructure as an evolving socio-technical system. Open systems theory, however, pushes the envelope of our understanding further by considering IT infrastructure design as a synthesis of differentiation and integration - adaptively configurable and reconfigurable over time. It provides a systematic platform to reduce a system to a set of subsystem assets, look at asset interactions, and holistically view the overall impact of this portfolio of subsystem assets. In addition, by juxtaposing concepts of homeostasis and goal-orientation with the dynamics of adaptive organization and configuration of IT infrastructure, open systems theory extends previous efforts with a systematic unraveling of the IT infrastructure puzzle. In short, open systems theory encapsulates the "inter-dependence, intricacy, and interweaving of people, systems, and processes are the culture bed of infrastructure" (Ciborra, 2001: 2). In the next section, we detail the elements of our proposed open-systems model of IT infrastructure.

IV. AN OPEN SYSTEMS VIEW OF IT INFRASTRUCTURE

The starting point for the model is the following: IT infrastructure is predominated by three major categories, best understood as modular subsystems: *content infrastructure* (e.g., databases, hard-drives), *computing infrastructure* (e.g., processors, monitors, programming tools), and/or *communication infrastructure* (e.g., routers, network operating systems). As modules of the IT infrastructure system, content, computing, and communication subsystems build on a combination of assets that are "structurally independent of each other, but work together…a framework…that allows for both independence of structure and integration of function" (Baldwin and Clark, 2000: 63).Three decades earlier, Robey (1977) voiced the need to understand IT infrastructure as a set of different components. Huber's (1984) seminal research was the first to accommodate this perspective, proposing a C^2 (computing and communication) view of IT infrastructure. The inclusion of databases expanded the view of IT infrastructure (e.g. Silver et al., 1995; Tapscott, 1997). These three modular categories, namely computing, content, and communications, provide a collectively exhaustive scheme encompassing an interrelated ontology of technology and people assets (as sub-subsystems) (Figure 2).

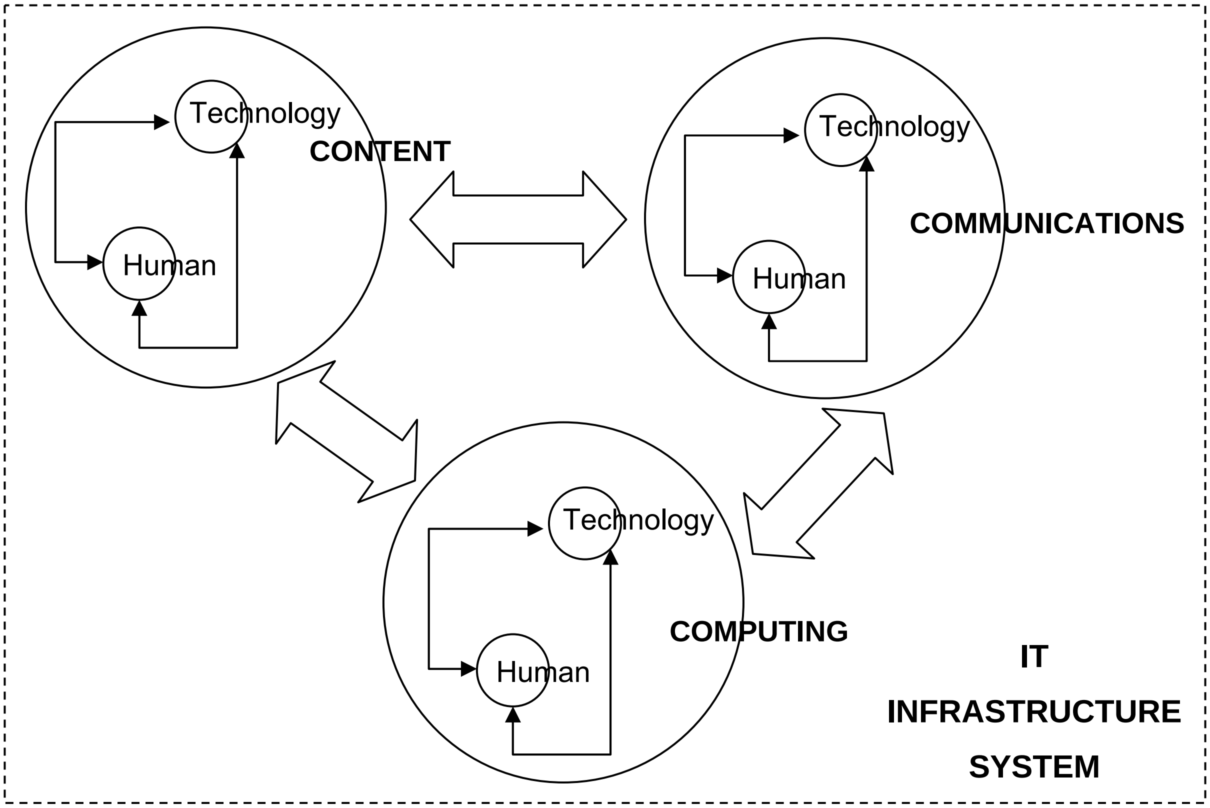


Figure 2: The three base modules of corporate IT infrastructure

The modular categorization of IT infrastructure alleges an IT infrastructure portfolio comprising of content, computing, and communication technologies at various levels of convergence (Kayworth et al., 1997; Sambamurthy and Zmud, 2000; Tapscott, 1997). Modular subsystem categories partition the infrastructure by levels of integration and convergence in underlying technology (operating and application) and human assets. This collectively exhaustive categorization accommodates a comprehensive IT infrastructure portfolio consisting of infrastructure resources at varying levels of convergence. The evolving organizational reality points towards a modular IT infrastructure portfolio configurable across all subsystem hierarchies.

TECHNOLOGY AND HUMAN ASSETS

IT infrastructure subsystems consist of technology and human assets. Technology assets are resources such as devices, applications, databases, and networks that are available as packaged components or as in-house customized developments. The technology subsystem is the installed base of IT infrastructure comprising of operating- and application-level assets. Operating-level assets are mainly hardware and device-specific technologies while application-level assets are program technologies that sit over the operating platform. Commonly, operating-level technology assets require proprietary adaptors to communicate directly with the hardware while application-level technology assets often rely on openly available adaptors (e.g., APIs) for indirect communications. Consequently, application-level technologies are relatively more flexible than operating-level technologies, given the proliferation of adaptors. Together, operating and application technologies help translate IT infrastructure as “a major business resource and a potential source for attaining sustainable competitive advantage” (Weill and Broadbent, 1998: 332), supporting an evolving service-driven design rather than a rigid artifact subject to inertia and incapable of adapting to demands posed by its environment.

Human assets complement technology assets. Human assets, by virtue of their “knowledge, skill sets, and experience.” serve as “the mortar that binds all technical IT components into robust and functional services” (Weill and Broadbent, 1998: 333). Human assets are the knowledge base of a corporate IT infrastructure. In an era where technology assets, by themselves, can no longer offer competitive advantages, it is only through the alignment and interaction of human assets and technology assets that the firms can capitalize on their corporate IT infrastructure. Nicholas Carr (2003), in his controversial article, notes that infrastructural technologies have become commodity inputs that, by themselves, do not offer any competitive benefits. Instead, IT infrastructure provides instrumental benefits only when paired or aligned with requisite human assets who serve as the minds behind the machine and create value through innovative uses of technology. Moving away from a deterministic view of technology as a necessary and sufficient representation of IT infrastructure, the strategic choice view advocates the importance of the recursive cycle of interaction and influence between human and technology assets, where human assets drive innovation out of commoditized technologies (e.g., Orlikowski, 1993; Ciborra et al., 2001). As Weill and Vitale (2002) remark, human IT infrastructure is the intelligence that is used to translate technology to create services for instrumental benefits.

MAJOR SUBSYSTEMS IN A COPRPORATE IT INFRASTRUCTURE PORTFOLIO

The corporate IT infrastructure portfolio has a set of subsystems organized as one or more of the three dominant subsystem classes: segmented, shared, and seamless. A segmented IT infrastructure class consists of content (A), computing (B), and communication (C) subsystems with little or no sharing of assets. The class of shared IT infrastructure consists of two subsystems sharing assets, e.g., content and computing (D), computing and communications (F). Finally, a seamless IT infrastructure class consists of integrated and tightly bound content, computing and communication assets (G). Figure 3 depicts the classes based on configurations of subsystem modules. The collectively exhaustive classes used to depict the IT infrastructure subsystem (Z) is shown in the Figure 3 where $A, B, C, D, E, F, G \subset Z$.

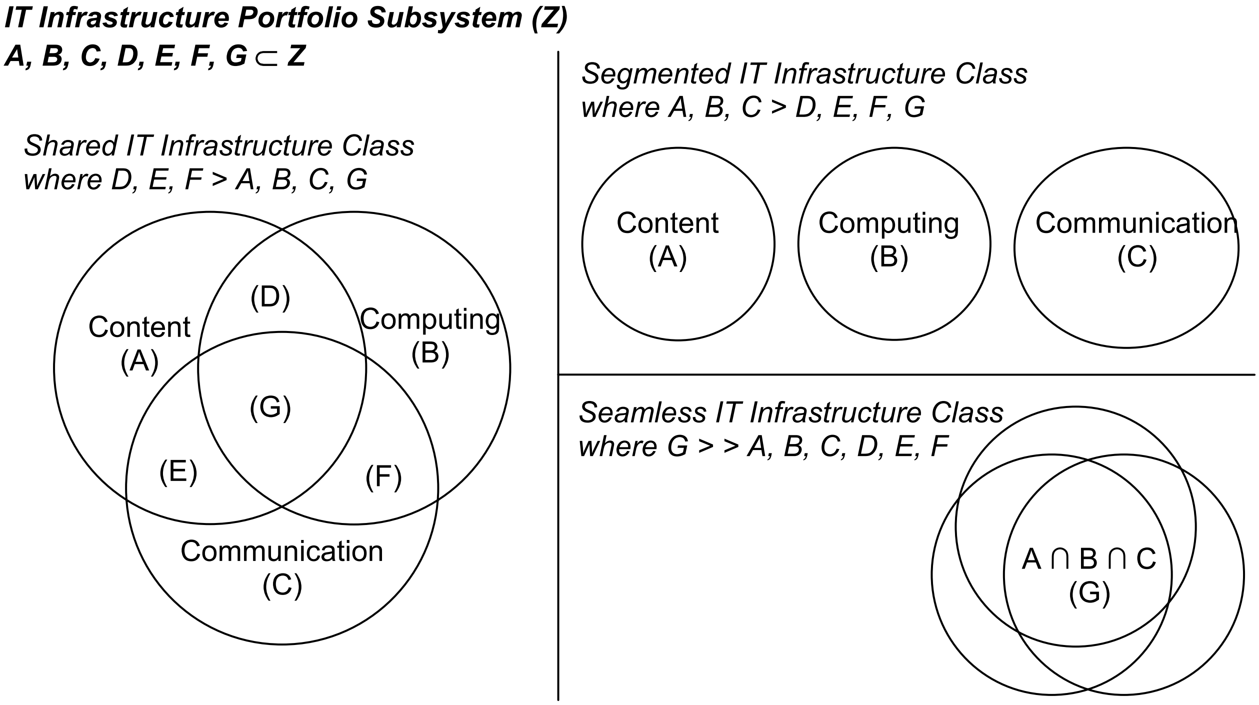


Figure 3: Modular Configuration Classes of an IT Infrastructure Subsystem

Any corporate IT infrastructure consists of most of these subsystem classes as a part of its portfolio: some segmented, some shared, and some seamless. A modular representation and design of IT infrastructure allows organizations to organize and configure their portfolio to meet service demands. Overall, based on different modular configurations of the IT infrastructure subsystem, corporate IT infrastructure can be organized as a portfolio comprising of different classes of configurations. The seven IT infrastructure subsystems occupy one of the three classes by virtue of their underlying technology and human assets, as shown in Table 1.

Table 1: Technology and Human Assets for IT Infrastructure Subsystems

Subsystems	IT Infrastructure Subsystem Assets		
	Technology Assets		Human Assets
	Operating Level	**Application Level**	**(trained in)**
Content	Disk drives, Storage Devices	Databases, Spreadsheets, Word Processors	Database Administration, Data Modeling
Computing	Hardware, Device upgrades	OS, Systems development	System Analysis, Software Engineering
Communications	Routers, Directory services	Network security, Network monitoring	System Administration, Network Design
Content/ Computing	Backup, Storage Systems	Content administration, Data mining	Business Intelligence, OLTP
Content/ Communications	SAN, NAS	Search tools, Distributed Storage	Web Development, Information Security
Computing/ Communications	Network OS, Thin Clients	Collaborative computing, Network Applications	Grid Computing, Application Security
Content/ Computing/ Communications	Enterprise servers	ERP suites, Collaborative tools	Enterprise Integration, Enterprise Security

CLASS I. SEGREGATED (FRAGMENTED) IT INFRASTRUCTURE SUBSYSTEMS

Content Subsystem (information-based assets) (A): The content subsystem includes all data and information-related assets governed by the corporate IT infrastructure. The content subsystem includes technology and human assets used for the acquisition, allocation and development of data and content resources needed to organize data for the purposes of cross-referencing and retrieval through the creation of information or data repositories (Keen, 1991).

Technology and Human Assets: Operating-level assets include magnetic-media storage (disk drives, external/removable storage devices), optical-media storage (CD, DVD); application-level assets focus on data creation and manipulation (databases, spreadsheets, text/graphic editors, statistical software). Human assets are resources such as database administrators, designers, and modelers used to develop, support and maintain content technologies.

Computing Subsystem (processors and system-based assets) (B): The computing subsystem includes processor-based resources focused on input-output, control, and processing, consisting of operating systems environments, system applications software, and technical standards for the hardware for operation and multi-vendor compatibility (Keen, 1991).

Technology and Human Assets: Operating-level assets include hardware such as processor-based systems (Sun, Unix, PC, Apple), mobile-devices (PDAs, pagers), input/output devices (keyboards, monitors, printers); application-level assets include stand-alone developmental software (compilers, debuggers, programming tools), system administration software (backup/recovery, emulators, system monitoring software, user management applications). Human assets for the computing subsystem include programmers, systems analysts, software engineers, testers, and systems maintenance personnel.

Communication Subsystem (network-based assets) (C): The communication subsystem deals with network-based resources used to support communications and provide organizational connectivity using voice and data networks, protocols, and standards (Keen, 1991).

Technology and Human Assets: Operating-level assets include physical hardware technologies (telephones, fax machines, routers), directory services (ADSI, X.500/LDAP), connectivity technologies (ATM. Gigabit Ethernet), network architecture (LAN, client/server, peer-to-peer); application-level assets include applications pertaining to network administration (network solutions, traffic management), network protocols (VoIP, DHCP, HTTP) and network troubleshooting. Human assets include personnel such as network administrators, network designers, telecommunication analysts, network service representatives, among others.

CLASS II. SHARED IT INFRASTRUCTURE SUBSYSTEMS

Content & Computing Subsystem (information & system-based assets) (D = A ∩ B)**:** The convergence or integration of content and computing assets gains significance especially in light of services such as data mining and business intelligence. Integration of computing and content assets relates to large scale processing of databases and application data, based on human and technology assets.

Technology and Human Assets: Operating-level assets primarily include computing (system) hardware resources that provide access to stored content such as separate backup and storage devices while application-level assets include applications pertaining to content administration, heterogeneous storage integration (data migration and synchronization) and content processing (data warehousing, data mining, data query processing). Human assets are people who can develop, support and manage integrated content and computing assets such as specialists in application data integration, OLTP, and data mining.

Computing & Communications Subsystem (system & network-based assets) (E = B ∩ C): The integration of system and network assets is evidenced by the growth of distributed large scale processing services where processing resources are being connected via popular network protocols. This infrastructure subsystem refers to technologies that address and help integrate computing (system processing) and communications (networks), typically high end computing clusters, by connecting processors and workstations over networks based on load distribution to optimize processes and resources (e.g., Sun UltraSPARC III based computing clusters).

Technology and Human Assets: Operating-level assets include technologies pertaining to secure systems-access, web applications, thin clients and terminals, network OS, distributed processing; application-level assets include distributed application performance monitoring, collaborative computing, heterogeneous system connectivity (CORBA, COM+/DCOM, middleware interoperability). In this subsystem, human assets include personnel trained in the operation, development, and maintenance of grid computing, application integration and security, clustering, middleware development among others.

Content & Communication Subsystem (information & network-based assets) (F = A ∩ C)**:** As information sources have become distributed over networked environments, the need for information integration has grown steadily (Rudensteiner et al., 2000). The Internet, particularly web-based developments have propelled the growth of integrating distributed content. Linking content repositories across the globe is becoming more and more popular, evidenced by the growth of the Internet and the World Wide Web along with enterprise search tools and networked content. What used to be a discrete, self-contained application on a server is being replaced by large relational databases at the back end and a flexible interface at the front end, connected by middleware (Coffin, 2006).

Technology and Human Assets: Operating-level assets include technologies for the preparation, deployment, and management of content over large networks (e.g., file and content servers, Network attached Storage, Storage Area Networks). Application-level assets include programs related to networked content security and assurance, search engines and interfaces and standards. Human assets are resources who develop, plan, and manage the integrated content and communication subsystem such as personnel involved in web-development, data security and assurance, server platform engineers, among others.

CLASS III. SEAMLESS IT INFRASTRUCTURE SUBSYSTEMS

Content, Computing, and Communication Assets (information, system, and network-based assets) (G = A ∩ B ∩ C)**:** The growth of enterprise systems and needs for enterprise-wide services has driven the integration of content, computing and communication subsystems, merging information, system and network-based assets. Enterprise application integration (EAI) is an example of that combination of processes, software, standards, and hardware resulting in the seamless integration of two or more enterprise systems allowing them to operate as one, such as building CRM systems, business-to-business integration, or leveraging legacy systems.

Technology and Human Assets: Operating-level assets include enterprise servers and enterprise storage systems for processing, hosting and serving information from and to distributed sources and recipients. Application-level assets include groupware, CRM, SCM, and ERP suites. Human assets support the development, installation, and maintenance related to content, computing and communication subsystem integration and include personnel trained in enterprise application integration, configuration management, ERP consultants, among others.

SERVICE-DRIVEN IT INFRASTRUCTURE DESIGN

A service-driven IT infrastructure design is demand-centric. Rather than deploying routine services from specific infrastructure subsystems, a service-driven IT infrastructure is designed to accommodate service requests. Central to this thesis is that generic service deployment is no longer value-added. Different firms require a portfolio of value-added services that can separate them from their competition. Correspondingly, firms need to maintain a portfolio of IT infrastructure resources that can be mapped to their portfolio of value-driven services. It is to be noted that a service-driven IT infrastructure design does not make generic services obsolete. Instead, service-driven IT infrastructure design allows firms to periodically reorganize its IT infrastructure design to emphasize on value-added IT services. According to Weill and Broadbent (1998), in a competitive environment, firms must be able to invent and reinvent IT infrastructure services to maintain a competitive edge. Therefore, facing a short shelf life for IT infrastructure services, firms need to constantly review and reorganize the underlying infrastructure to fulfill service demands.

Underpinning the design and reorganization (reconfiguration) of content, computing, and communication subsystems is the reconfiguration of underlying technology and human assets for reengineering processes to deliver service. Reconfiguring technology and human assets drives the reorganization of the IT infrastructure. For example, the need for Web-based services (content and communications) drives the employment of web-development personnel while the outsourcing of data centers would reduce the number of database personnel. Similarly, service demands drive the integration, modularization, upgrading, substitution and outmoding of technologies. For example, standardization and content availability services may prompt companies to consolidate disparate databases into storage area networks and reduce staffing (*Computerworld*, 2006). Likewise, demand for specific IT services lead to reorganization, reconfiguration and reengineering of IT infrastructure processes. A common example is how enterprise integration services lead to the reconfiguration and reengineering of common business processes and require a class of seamless assets for positive service outcomes.

In a service-driven IT infrastructure, the degree of modularity/integration of subsystems is a fundamental design consideration, with major consequences in terms of cost, flexibility, efficiency and functionality. Firms have a wide range of options available to them, from designs offering a high level of modularity and flexibility such as “modular” web services to enterprise-centric architectures and highly integrated solutions such as ERP systems (Hagel and Brown, 2001). Although a firm’s IT infrastructure design is a spillover of many decisions made over time - comprising of legacy systems as well as more recent additions, the ability of a firm to design its IT infrastructure to adapt to service demands from its environment is key to its sustenance. As a result, a service-driven IT infrastructure tends to be more modular, allowing for rapid reconfigurations.

Different service orientations require different levels of IT infrastructure design decisions. A firm that emphasizes more on remote access services may reorganize its IT infrastructure to focus more on its communications resources while outsourcing its data center (content) and maintenance (computing) services (as managed services). A company emphasizing on collaborative computing services (e.g., grid computing) can focus more on integrating its computing and communication resources. A company enterprise emphasizing on locally available enterprise services can focus more on integrating their content while a company emphasizing on global availability of enterprise services would focus more on integrating their content and communication infrastructure resources. Because a firm relies on delivering a portfolio of services, some outsourced (managed) and some in-house, their IT infrastructure needs are different, leading to different IT infrastructure designs. As new services are demanded by evolving strategies, firms have to reorganize, upgrade, and scale their IT infrastructure resources to accommodate these service changes. For example, Honda Motor Corporation’s strategy to reduce cycle time in new car production (see Broadbent, 1995) required a set of collaborative services. The move would require Honda to reorganize its IT infrastructure design to focus on integrating its content and communication resources so that designers could have access to content (e.g., drawing specs, fabrication templates) in real-time.

The ability to reorganize the IT infrastructure portfolio based on services that a firm decides to deliver is a key to sustenance, especially during a time when IT planning is supporting a set of core services in-house and outsourcing other services. In the face of realities, from outsourcing to mergers and acquisitions, a company’s service portfolio is constantly changing, therefore requiring a reengineering of processes and a corresponding reorganization of the underlying IT infrastructure. Two factors, namely synergistic specificity and heterogeneity of service demands, seem to influence the organization of an IT infrastructure for alignment or fit.

First, *synergistic specificity* is “the degree to which a system achieves greater functionality by its components being specific to one another” (Schilling, 2000: 316). Synergistic specificity creates a uniquely interdependent condition that can sometimes become a constraint. For example, certain IT infrastructure portfolio designs may create lock-in effects. For instance, a company that has deployed an ERP system may experience a lock-in by tying its services too tightly with its integrated assets base. It is not uncommon to hear how firms find it exceedingly difficult to reduce dependence on an ERP consultant or a set of processes mandated by the technology. Similarly, a firm’s IT infrastructure may become captive to segmented assets due to difficulties of integration (e.g., lack of available adaptors, integration points). In short, if a firm's IT infrastructure is too tightly coupled with a particular service portfolio, the unique interdependence may create a degree of specificity and therefore create a degree of inertia in its IT infrastructure design.

Second, *heterogeneity of service demands* refers to the varied levels of service requests. A corporate IT infrastructure, as an open system, is constantly influenced by the service demands placed upon it by the environment,, e.g., the market. A diversity of markets faced by a firm increases the heterogeneity. Some firms, for example, are a collection of disparate business divisions, with different products, sales channels, etc., and may even be in different industries. In these situations, a key advantage of modular IT infrastructure components is that they can be combined in different ways to meet the unique requirements of individual business/functional units. Conversely, economies of scope can be exploited in situations where there are common needs across the business/functional units, with benefits accruing to a firm through the sharing of integrated IT infrastructure resources (Sambamurthy and Zmud, 1999). Heterogeneity of inputs and demands will have indirect effects, reinforcing the pressure created by the other (Schilling, 2000). The availability of modular IT infrastructure subsystem assets, combined with demands for differentiated IT services and applications within the firm, can possibly influence the migration of the IT infrastructure between segmentation and seamless integration.

What aligns the design of the IT infrastructure portfolio to fulfill or deliver a set of services is the notion of fitness – the degree to which subsystems and the environment are "mutually acceptable" (Alexander, 1964: 19). The assumption is that systems respond to changes in the environment or changes in their components through adaptation or evolution, seeking better fitness (Holland, 1999; Schilling, 2000). IT infrastructure modularity reflects these fundamental ideas, with migration towards (or away from) increasing modularity as the adaptive response to service demands arising from changes in the environment, both internal and external to the firm.

A TYPOLOGY OF IT INFRASTRUCTURE SERVICES

Any service driven IT infrastructure portfolio aims at fulfilling a set of services that reduce the gap between purported an actual services. An IT infrastructure driven by services must offer a predesignated set of services that the company maintains as its competitive edge. For example, a company like Wipro that maintains its competitive edge through backend operations and maintenance support will have align its IT infrastructure portfolio to meet its service demands, thus focusing on a more shared communication and content IT infrastructure. On the other hand, a company such as Pixar whose competitive edge is by developing large-scale complex animations may focus more on its computing infrastructure to align it to its services.

To deliver a set of services, a firm must choose successfully choose across four service dimensions. transactional services, operational efficiency services, operational quality services, and strategic services. Transactional services are short-term services that reduce costs of transactions (Weill and Olson 1989). A transaction service may be a login service, a point-of-sale service, or a computing/ processing service. Operational efficiency services are more medium-term services that aim to optimize output for a reduced set of inputs. Operational efficiency services include material scheduling and Kanban services, batch setup services, or query optimization services. Operational quality services are, on the other hand, while also medium-term services, less focused on efficiencies and more on the effectiveness of operations. At Google, search, mapping, and email hosting appliances to process and handle queries swiftly and efficiently form its transactional services core. At Infosys or Wipro, routine scheduling and maintenance operations based on elements such as resource consumption and utilization per hour make up their operational efficiency core. At SAS, business intelligence and data mining services aimed at knowledge discovery and information sharing constitute operational quality services. And the innovation, development, and deployment of client based solutions that allow Accenture or IBM Global Services to increase market share or offer competitive differentiation for their clients is a form of strategic services.

To add perspective, Table 2 lists a typical organizational service portfolio, tying them to corresponding IT infrastructure subsystem assets. Interestingly, as services are outsourced, so are corresponding IT infrastructure assets. For example, if a firm decides to outsource its firmwide messaging services as a managed service, it will require fewer IT infrastructure assets integrating content and communication resources. Consequently, the firm may decide to undo its email servers (technology assets) and downsize personnel related to developing solutions for (e.g., scripts) and maintaining the email server. Take for example, Amazon.com's spirited debut into the world of managed hosting. In a recent USA Today column by Maney (2006), Jeff Bezos promotes Amazon.com's hosting services as a useful alternative for firms to divest their infrastructure and therefore reduce worries stemming from ongoing efforts of control and monitoring. Similar alternatives available from outsourcing and hosting point are just a harbinger of how corporate IT infrastructures evolve over time. A service-oriented IT Infrastructure system migrates by reorganizing its portfolio of assets in perspective of competitive and market demands: placing more emphasis on some subsystems and reducing emphasis on the others. Through the close interplay of technology and human assets, firms are constantly evolving by balancing their IT infrastructure portfolio to align it with their portfolio of services as demanded by their proximal environment. Feedbacks from service outcomes of an IT infrastructure portfolio are used to reorganize subsystem assets within the IT infrastructure. It is a phenomenon of continuous improvement.

Table 2. Services and Corresponding IT Infrastructure Requirements in an Organizational Portfolio.

Services *(adapted from and extending the list of services by Broadbent and Weill (1997))*	**Infrastructure Subsystems**						
	A	**B**	**C**	**D**	**E**	**F**	**G**
Firmwide Communication Services			√				
Firmwide Messaging Services (e.g., Email)					√		
Standard Setting Services				√	√	√	√
Data Security & Assurance (incl. Backups)	√			√	√		
Data Center Services				√			
Desktop Application Services		√					
Business Intelligence	√			√			
Shared Backup Services					√		
Remote Access Services (Mobile)					√	√	
Call Center Services					√		
IS Project Management Services							√
Enterprise Computing						√	
Value-Added Network (VAN) Services					√		
Enterprise Services (ERP, SCM)							√
Workstation Network Services						√	
Online services (online apps & content)					√	√	
Network Security			√				
Outsourced Services							

Legend: A: Content; B: Computing; C: Communications; D: Content & Computing; E: Content & Communications; F: Computing & Communications; G: Content & Computing & Communications.

A PORTFOLIO MODEL OF IT INFRASTRUCTURE DESIGN

As aforementioned, each IT infrastructure subsystem with its technology and human assets contributes towards delivering value in varying degrees. The portfolio of IT infrastructure subsystems must be aligned with the services delivery portfolio requirements. We agree with Ciborra et al.'s (2001) recommendation that the IT infrastructure portfolio must be managed like an investment portfolio consisting of modular investment assets (various financial instruments). The portfolio must be reorganized and reconfigured to balance the weights of assets to suit particular service outcomes. Service-driven IT infrastructure design decisions can therefore be formulated as a portfolio selection problem for maximizing service returns and minimizing the entropic gap, where the expected return is the return from all allocated subsystem assets in an organization's IT infrastructure portfolio. In optimizing the portfolio by reducing the entropic gap from the difference between expected and actual service returns from a given IT infrastructure portfolio made up of different allocation of subsystem assets, organizations can reconfigure their IT infrastructure subsystem assets

Assume that a firm has an IT infrastructure M_0 configurable into n (n=7, i.e., A, B,….G) possible subsystem assets, S_j, j=1, 2, …, n. Let R_j be the set of services offered by the subsystem S_j (a random variable). The level of presence of each of the seven subsystem assets in an IT infrastructure portfolio can be understood by its allocated weight, x_j, from M_0 towards subsystems S_j where $x_j \geq 0$ and $\sum_{j=1} x_j = M_0$.

Let $E(R)$ denote the mathematical expectation of a random variable R to define expected service returns (output) from a particular IT infrastructure subsystem:

$$r_j = E(R_j),\ q_j = E\ (E(R_j) - R)$$

where r_j and q_j are the expected rate of service returns and the expected entropic gap (deviation between expected and actual service returns) for asset S_j, respectively.

Therefore, the expected service returns of an IT infrastructure portfolio $\boldsymbol{x} = (x_1, \ldots., x_n)$ is shown by

$$R(x_i, \ldots., x_n) = E\ [\textstyle\sum_{j=1} R_j x_j] = \sum_{j=1} E(R_j) x_j = \sum_{j=1} r_j x_j$$

We use Feinstein and Thapa (1993) and Cai et al. (2000) portfolio selection model where

$x_j \geq 0$, the maximum tolerable entropic gap for service returns can be restated as a risk function defined as,

$$w(x) = \max_{1 \leq j \leq n} E\,(E(R_j) - R)x_j = \max_{1 \leq j \leq n} q_j x_j$$

An organization, in designing its IT infrastructure portfolio, typically has an acceptable level of entropic gap that it can tolerate in its delivery of services. Assuming that the investor's weight on risk is λ, lying within the specific interval [0, 1], an organization can optimize its IT infrastructure portfolio ***ITPO**(p))* as follows,

$$\text{minimize } \lambda \max_{1 \leq j \leq n} q_j x_j - (1-\lambda) \textstyle\sum_{j=1} r_j x_j$$
$$\text{subject to } \textstyle\sum_{j=1} x_j = M_0.$$
$$x_j \geq 0,\ j = 1, \ldots., n$$

assuming that, $q_j > 0$, $j = 1, \ldots., n$ (i.e., all service returns from subsystem assets exhibit some entropic gap) and where the degree of reconfiguration of IT infrastructure portfolio φ_{ITR} is subject only when $R_j - E(R_j) < 0$ and not otherwise (i.e. 0).

The aforementioned optimization model can be used to asses the performance of any particular service-driven IT infrastructure portfolio to evaluate their service returns in a way to minimize the entropic gap by reorganizing (rebalancing) the IT portfolio mix of subsystem assets. By migrating levels of integration across modular subsystems, an organization may be able to periodically assign different weights to its IT infrastructure assets to achieve a degree of alignment or fit with its service portfolio.

V. DISCUSSION

From 1980 to 2000, approximately 50 percent of capital investments were in information technology, directed primarily towards building, maintaining and upgrading IT infrastructure (Westland and Clark, 1999). Even with recent budget cutbacks, global IT expenditures are currently about $1 trillion annually (Carr, 2005). At the same time, approaches to IT decision-making have become more sophisticated, moving beyond ROI analysis of individual projects and infrastructure component purchases to more macro approaches. Valuing and prioritizing portfolios of IT projects and consideration of infrastructure investments within the context of the overall IT architectural design have become commonplace (Bardhan et al., 2004; Ardagna and Francalanci, 2005). Consequently, the development of guidance for decision-making related to IT investments depends upon framings of the macro-level artifact of IT infrastructure as well as its constituent parts. Similarly, the development of theory relating specific attributes or capabilities of a firm's IT infrastructure (e.g., flexibility, "reach and range") to organizational outcomes (e.g., business process redesign success, innovation) requires conceptualizations at this level (Henderson and Venkatraman, 1994; Broadbent et al., 1999; Byrd and Turner, 2000; Kumar, 2004).

One of the most challenging issues facing IT infrastructure designers – those who design products for the marketplace and those responsible for architecting their firms' IT infrastructures – is a simple yet comprehensive account of corporate IT infrastructure. Similarly, researchers (e.g. Ciborra et al., 2001; Edwards, 2003) have contended that IT infrastructure is perplexing because of a lack of good models to apply to its analysis. Answering the challenge, the paper uses concepts from open systems theory to forward an organizing logic for corporate IT infrastructure design as a modular portfolio of assets - reconfigurable, open, and service-driven: evolving and adapting to changes in its environment.

The primary merit of the open systems perspective is its ability to offer a simple yet comprehensive texture of an evolving IT infrastructure. The perspective balances both synthesis and decomposition. Systems concepts contribute to our ability to view IT infrastructure as a configurable system. IT planning budgets and investments towards creating and maintaining an IT infrastructure serve as input to the system. The input serves to pay for personnel and technology that can be combined to create a modular IT infrastructure that offers dynamic configuration. The expected output is a modular and configurable IT infrastructure at the level needed by the particular firm. Table 3 synthesizes key systems concepts and their corresponding exemplification in IT infrastructure design.

Table 3. Synthesis of Systems Concepts

Term	Definition	Exemplification
Receptor	Service demands and resources flowing into a system/subsystem	IT infrastructure investments and service-oriented triggers
Controller	System processes by which inputs are transformed into outputs	Combining and recombining IT technical infrastructure modules and HR to create a set of specific and shared services for the organization

Effector	Service outcomes resulting from the IT infrastructure system throughput	Modularity in IT infrastructure design and the shared services portfolio
Entropic Gap	The margin between actual and expected output from a given set of inputs an the existing throughput	The difference between the expected and realized IT infrastructure design and shared services portfolio
Environment	Forces outside a system or subsystem that exert triggers for adaptation and change	Market and organizational forces triggering different IT infrastructure choices and the migration to/from modular IT infrastructure designs
Modularity	The ability to decompose, combine, and recombine a system and its subsystems into independent constituents	Combining the recombining IT infrastructure modules and HR to deliver the requisite organizational services
Feedback	The dynamic ability to recombine a system and its subsystems based on a given output.	Recombining IT infrastructure components to migrate towards or away from segmentation or seamless designs based on the existing IT infrastructure design and services portfolio

To recapitulate, the gist of corporate IT infrastructure as an open system is as follows: A corporate IT infrastructure is an open system. Because a firm operates within and caters to one or more markets, the market places service demands on the firm. Receptors scan the market environment and periodically receive service requests (e.g., need to provide enterprise services, need to offer remote access to premium partners). The portfolio of service requests is relayed to the controller, i.e., the IT infrastructure system, which adapts based on an optimal (re)configuration of IT infrastructure subsystem assets to deliver the required services. The subsystems are self-contained (i.e., they serve stand-alone functions) and cohesive in that they can, although not always, be made to communicate between each other using open standards (e.g., HTML, XML), protocols or application programming interfaces (APIs). The functionality and operational parameters of each subsystem define its boundary.

Upon identifying required services (including opportunities for outsourced services), the controller, i.e. the infrastructure planning/management function in the firm, organizes its IT infrastructure assets correspondingly (e.g., through upgrades, hires, downsizing, cross-training, integration, segmentation). Once the IT infrastructure portfolio has been organized to meet service demands, the effector reviews the resulting services from the particular IT infrastructure portfolio configuration by comparing the actual service outputs vis-à-vis service expectations. The objective of organizing the IT infrastructure subsystem assets is to minimize the entropic gap. If actual service outputs surpass expectations, the effector provides positive feedback to the controller; if actual services fall short of expectations (depending on the degree of tolerance for deviations), the effector provides negative feedback to the controller. The controller, in turn, uses the feedback to reconfigure or reorganize its IT infrastructure subsystem assets. The process is continuous, aimed at adapting to changes in service demands from the environment and reconfiguring IT infrastructure subsystem assets to maintain homeostasis. Through modular configurations and reconfigurations, an IT infrastructure remains service driven, migrating and evolving to respond to its environment. For example, in the nineties, the market trend from mainframe toward client-server computing led to many companies to revamp their IT infrastructure by configuring mainframes as surrogate file servers in a client-server arrangement. By continuously reconfiguring its technology portfolio, the goal of the IT infrastructure, as a system, is to create a verifiable and value-added infrastructure that is scalable and aligned with business needs. The open systems lens thus offers an overarching perspective of IT infrastructure as a dynamic, modular, and configurable IT artifact (Figure 4).

At this point, it is important to note that the aforementioned overview is an ideal and rational synthesis of corporate IT infrastructure design and can thus be somewhat limiting. In reality, corporate IT infrastructure designs are far from rational and exhibit various degrees of dysfunctionalities, stemming from a variety of issues: push selling by vendors, lock-in effects, deployment problems, personnel turnover and attrition, governance, among many others. However, we argue that dysfunctionalities are often manifested in the entropic gaps in service, leading to requisite feedback to reorganize the system.

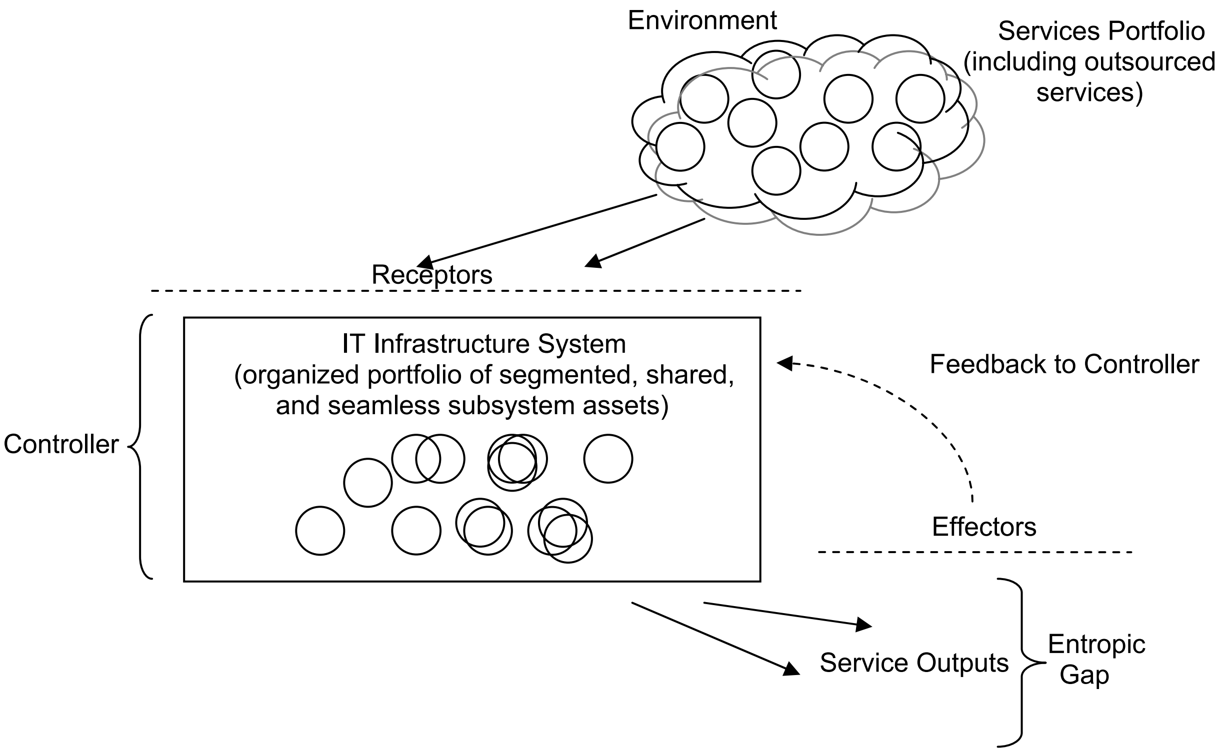


Figure 4: An Open Systems Perspective of Corporate IT Infrastructure Design

VI. CONCLUSION

The paper adds to our understanding of IT infrastructure in several ways. We have proposed a systems model of IT infrastructure that addresses important drawbacks of prior conceptualizations. Our re-conceptualization of IT infrastructure has implications for both practice and research in terms of the new understandings it provides and the opportunities for investigation that it affords IS researchers. This departure from prior conceptualizations offers a rich underpinning for future investigations and theory development. We believe that the proposed IT infrastructure model represents a major step in addressing the call for reconceptualizing IT artifacts from being “multiple, fragmented, partial, and provisional,” to developing models capable of supporting theoretical work that reflects “the emergence and evolution of IT artifacts as complex and changing technosocial processes existing in time and over time” (Orlikowski and Iacono, 2001: 132).

First, an immediate contribution to practice is the availability of a new representation of IT infrastructure that provides an analytical tool for decision-makers and infrastructure designers. For firms in the process of developing or updating their IT infrastructure, the ability to classify their adopted/proposed IT portfolio using the systems model we have described should prove valuable in representing and evaluating different options. If firms can logically separate technologies by their functionality, human resources and services, resource allocation choices and the implications of those choices can become more apparent. A systemic view of IT infrastructure allows firms to unbundle their IT infrastructure into observable and classifiable components. This allows firms to better inventory and allocate their technology portfolio to align and serve organizational goals. For example, while there are strong pressures from the market towards integrated assets, they have not supplanted modular technologies. Each requires different HR support and provides different services. It is imperative that firms consider these technologies separately to understand their fit and alignment with organizational directions. While any organizational IT infrastructure can be classified by its portfolio of segmented, shared, and seamless infrastructure subsystem assets, portfolio weights vary across firms, each with different alignment considerations driven by business needs. For example, a firm that recognizes that its infrastructure needs are autonomous may advise its strategic business units to adopt modular technologies rather than look at integrated solutions. Conversely, some entrants and incumbents may feel strong pressures for adopting relatively more integrated IT infrastructure assets. For instance, firms offering e-business services or engaged in distributed manufacturing have a greater need for enterprise-wide information sharing. Because modular technologies sometimes suffer from standards incompatibility, the choice of shared and seamless classes of IT infrastructure solutions may be more appropriate in their aligning of IT infrastructure with business needs. Even while choosing integrated systems such as ERP applications, firms prefer flexibility. Consequently, ERP systems are being decomposed and adopted in modules (financial-control, materials, manufacturing) so as to maintain flexibility and configurability. It is only by re-examining IT infrastructure as a system that firms can capture the significance of their choice and use of technologies under simultaneous considerations of modularity and integration.

For researchers, use of the systems paradigm to understand IT infrastructure offers new horizons for investigation. First, it removes the traditional bundling of all technologies into one basket, equating databases with servers. Instead, it offers a logical classification based on functionality. Furthermore, the systems lens allows us to capture the dynamics of an IT infrastructure. Not only can we unbundle IT infrastructure technologies, but we can also represent the modular variations through configurations and reconfigurations over time. Consequently, we are able to acknowledge the entire IT infrastructure portfolio as a mix of technology and human assets, including segmented, partly integrated, and seamlessly integrated infrastructure assets. In addition, we can trace the corresponding human resources infrastructure and link them to the shared set of services offered. By using systems theory to derive a specific conceptualization of IT infrastructure, this paper has demonstrated the scope offered therein. The principles of general systems theory can be utilized with our framework to investigate and derive other explanatory models surrounding this IT artifact and other organizational systems. For instance, what triggers particular choices of infrastructure technologies? What factors determine whether the organization or the market serves as the salient environment in shaping these choices? How does the environment impact the IT infrastructure systems and determine its choice of technology adoption? Why do particular firms tend to adopt more modular technologies than others? How do firms align their IT infrastructure and dynamically reconfigure their infrastructure to match environmental demands? How is entropy measured before a move towards realignment and reconfiguration? Are there major operational challenges and how can they be resolved? These questions suggest a number of avenues for future research. In addition to further conceptual work, priorities for research in the near term include the empirical investigation of the proposed model to examine how services drive corporate IT infrastructure reorganization. Another potential opportunity would be to integrate the systems model of IT infrastructure with Kumar's (2004) framework for assessing the business value of IT infrastructures, loosely constructed on the assumption of IT infrastructure as a dynamic asset.

In summary, the systems perspective allows for a deeper and more cognizant understanding of IT infrastructure as an organizational artifact. The conceptual breadth and scope offered by systems theory opens up new avenues for investigation in the near future. This research is an attempt to introduce its richness of scope and exemplify its use in conceptualizing IT artifacts. We move away from an isolationist perspective that overly simplifies the complexities forwarded by reality (Cummings, 1980) to understand convergence and interaction. By doing so, we see how systems theory helps weave a fabric of rich complexity, yet maintaining a synthesis of purpose. The systems metaphor thus provides a comprehensive understanding of IT infrastructure at differing levels of abstraction, offering a fresh paradigm on which to build. If this theoretical lens spurs further development of richer models of IT artifacts that better capture reality, it will be a useful paradigm shift – one that will enable IS researchers to move forward in theorizing this critical IT artifact.

VI. REFERENCES

Alexander, D (1964) *Notes on the Synthesis of Form*. Harvard University Press, Baltimore.

Applegate, LM., Austin, A, and McFarlan, FW. *Corporate Information Strategy and Management: Text and Cases*. 7th ed. New York: Irwin/McGraw Hill, 2006.

Ardagna, D and Francalanci, C (2005) A cost-oriented approach for the design of IT architectures. *Journal of Information Technology* 20(1), 32-51.

Ashby, R (1956) *An Introduction to Cybernetics*. Chapman & Hall, London.

Baldwin, CY and Clark, KB, (2000) *Design Rules*, MIT Press.

Bardhan, I, Bagchi, S and Sougstad, R (2004) Prioritizing a portfolio of information technology investment projects. *Journal of Management Information Systems* 21(2), 33-60.

Bertalanffy, L Von (1968) *General Systems Theory: Foundations, Development, Applications*. George Braziller, New York.

Boulding, KE (1956) General systems theory – the skeleton of science. *Management Science* 2(3), 197-208.

Bowker, GC and Star, SL (1999) *Sorting Things Out: Classification and its Consequences*, MIT Press, Cambridge, MA.

Broadbent, M (1995) *Honda's Civic Lesson*, Business Week, September 18, 26-28.

Broadbent, M., and Weill, P. (1997) *Management by Maxim: How Business and IT Managers can Create IT Infrastructures*. *Sloan Management Review* 38, 3. Spring, 77-92. Reprint No. 3836

Broadbent, M, Weill, P, O'Brien, T and Neo, BS (1996) Firm context and patterns of IT infrastructure capability. In *Proceedings of the Seventeenth International Conference on Information Systems* (Degross J, Jarvenpaa SL and Srinivansan, A, Eds.) pp. 174-194.

Broadbent, M, Weil, lP and St.Clair, D (1999) The implications of information technology infrastructure for business process redesign. *MIS Quarterly* 23(2), 159-182.

Byrd, TA and Turner, DE (2000) Measuring the flexibility of information technology infrastructure: exploratory analysis of a construct. *Journal of Management Information Systems* 17(1)167-208.

Cai, XQ, Teo, KL, Yang, XQ, and Zhou, XY, (2000) Portfolio optimization under a minimax rule. *Management Science* 46, 957-972.

Carr, NG (2005) The next IT revolution. *eWeek* 22(21), 44.

Ciborra, CU, Braa, K, Cordella, A, Dahlbom, B, Failla, A, Hanseth, O, Hepso, V, Ljungberg, J, Monteiro, E, and Simon, KA (eds.) (2001) *From Control to Drift. The Dynamics of Corporate Information Infrastructures*. Oxford University Press, New York.

Coffin, C (2006), As Networks and Applications Converge, New Challenges Emerge, Analyst Corner, *CIO*, February 28th.

Computerworld (2006) SNW: Users hope storage consolidation will lower TCO, November 02.

Cummings, TG (1980) *Systems Theory for Organizational Development*. John Wiley and Sons, New York.

Duncan, NB (1995) Capturing flexibility of information technology infrastructure: a study of resource characteristics and their measure. *Journal of Management Information Systems* 12(2), 37-57.

Edwards, PN (2003) Infrastructure and modernity: force, time, and social organization in the history of sociotechnical systems, in: *Modernity and Technology,* T.J. Misa, P. Brey and A. Feenberg (eds.), MIT Press, Cambridge, MA, pp. 185-226.

Emery, FE and Trist, EL (1965) The causal texture of organizational environments. *Human Relations* 18(1), 21-32.

Feinstein, CD and Thapa, MN (1993) A Reformulation of a Mean-Absolute Deviation Portfolio Optimization Model. *Management Science* 39, 1552-1553

Fomin, VV (2003) Standard Making: A Critical Research Frontier for Information Systems, *MISQ Special Issue Workshop.*

Hagel III, J and Brown, JS (2001) Your next IT strategy. *Harvard Business Review* 79(9), 105-113.

Hanseth, O and Lyytinen, K Theorizing about the design of information infrastructures: design kernel theories and principles, Working Paper

Henderson, JC, and Venkatraman, N (1994) Strategic alignment: A model for organizational transformation via information technology. In *Information Technology and the Corporation of the 1990's* (Allen, TJ and Scott Morton, MS, Eds) pp. 202-220, Oxford University Press, Oxford, UK.

Holland, J (1999) *Emergence: From Chaos to Order*. Perseus, Cambridge, MA.

Huber, G (1984) The nature and design of post-industrial organizations. *Management Science* 30(8), 928-951.

Kayworth, T, Sambamurthy, V and Chatterjee, D (1997) A conceptual framework of IT infrastructure: the critical link of technology standards. In *Proceedings of the Third Americas Conference on Information Systems* (Perkins, B and Vessey, I, Eds).

Keen P (1991) *Every Manager's Guide to Information Technology*. Harvard Business School Press, Boston, MA.

Kumar RL (2004) A framework for assessing the business value of information technology infrastructures. *Journal of Management Information Systems* 21(2), 11-32.

Lawrence, P and Lorsch, J (1967) *Organization and Environment: Managing Differentiation and Integration*. Harvard University Press, Boston, MA.

Lee DMS, Trauth, E and Farwell, D (1995) Critical skills and knowledge requirements of IS professionals: a joint academic/industry investigation. *MIS Quarterly* 19(3), 313-340.

Lewis BR and Byrd, TA (2003) Development of a measure for the information technology infrastructure construct. *European Journal of Information Systems* 12(2), 93-109.

Monteiro, E (2001) in Ciborra, CU, Braa, K, Cordella, A, Dahlbom, B, Failla, A, Hanseth, O, Hepso, V, Ljungberg, J, Monteiro, E, and Simon, KA (eds.) (2001) *From Control to Drift. The Dynamics of Corporate Information Infrastructures*. Oxford University Press, New York.

Orlikowski, WJ (1993) CASE Tools as organizational change: investigating iIncremental and radical changes in systems development. *MIS Quarterly*, 17(3), 309-340.

Orlikowski, WJ and Iacono, CS (2001) Research commentary: desperately seeking the 'IT' in IT research--a call to theorizing the IT artifact. *Information Systems Research* 12(2), 121-134.

Robey, D (1977) Computers and management structure: some empirical findings re-examined. *Human Relations* 30(11), 963-976.

Rudensteiner, E., Koller, A and Zhang, X (2000) Maintaining data warehouses over changing information sources. *Communications of the ACM* 43(6), 57-62.

Sambamurthy, V and Zmud, RW.(1999) Arrangements for information technology governance: a theory of multiple contingencies. *MIS Quarterly* 23(2), 261-290.

Sambamurthy, V and Zmud, RW (2000) Research commentary: the organizing logic for an enterprise's IT activities in the digital era: a prognosis of practice and a call for research. *Information Systems Research* 11(2), 105-114.

Schilling, M (2000) Towards a general modular systems theory and its application to inter-firm product modularity. *Academy of Management Review* 25(2), 312-324.

Scott, WG (1961) Organization theory: an overview and an appraisal. *Academy of Management Journal* 4(1), 7-26.

Scott, WR (1998) *Organizations: Rational, Natural, and Open Systems.* Prentice-Hall, Upper Saddle River, NJ.

Silver, MS, Markus, M, and Beath, C (1995) The information technology interaction model: a foundation for the MBA core course. *MIS Quarterly* 19(3), 361-390.

Simon, H (1981) *The Sciences of the Artificial*, 2nd edition. MIT Press, Boston, MA.

Skyttner, L (2001) *General Systems Theory: Ideas and Applications*, World Scientific: NJ.

Star, SL (1999) The ethnography of infrastructure. *American Behavioral Scientist* 43(3), 377-392.

Tapscott, D (1997) *The Digital Economy: Promise and Peril in the Age of Networked Intelligence*. McGraw-Hill Publication, New York.

Maney, K (2006) Amazon's new direction: Point, click, make a product to sell to the world, *USA Today*, November 21.

Van Gigch, JP and Kramer, NJTA (1981) A taxonomy of systems science. *International Journal of Man-Machine Studies* 14(2), 179-191.

Weill, P and Broadbent, M (1998) *Leveraging the New Infrastructure – How Market Leaders Capitalize on Information Technology*. Harvard Business School Press, Boston, MA.

Weill, P and Vitale, M (2002) What IT infrastructure capabilities are needed to implement e-business models? *MIS Quarterly Executive* 1(1), 17-34.

Westland JC and Clark THK (1999) *Global Electronic Commerce: Theory and Case Studies.* MIT Press: Cambridge.

Zachmann, JA (1987) A framework for information systems architecture. *IBM Systems Journal* 26(3), 276-292.