

Journal of Information Technology Management

ISSN #1042-1319

A Publication of the Association of Management

CLOUD COMPUTING: A NEW PHASE IN INFORMATION TECHNOLOGY MANAGEMENT

AL BENTO UNIVERSITY OF BALTIMORE abento@ubalt.edu

REGINA BENTO UNIVERSITY OF BALTIMORE rbento@ubalt.edu

ABSTRACT

The use of cloud computing is becoming widespread, but systematic study of its managerial implications is lacking. This paper examines cloud computing in the context of other major changes in Information Technology (IT) and explores the revolutionary transformations and challenges it brings to IT management. The paper analyzes the IT pendulum of centralization and decentralization and discusses the managerial implications of the major components of cloud computing: hardware (INTEL, IBM chips to support virtualization); services (Google, Amazon.com); applications (SaaS, Software-as-a-Service) and virtualization (VMware), or their combination (Citrix).

Keywords: cloud computing, Information Technology Management, centralization, decentralization, hardware, services, applications, virtualization, SaaS, VMware, Citrix

INTRODUCTION

The use of cloud computing is rapidly growing, and so is the literature on the technical issues of implementation. Our knowledge of the managerial implications of cloud computing, however, still lags far behind. This paper examines the phenomenon of cloud computing, places it in the context of other major changes in Information Technology (IT) and explores the potentially revolutionary transformations and challenges it brings to management.

The paper starts by analyzing the IT pendulum of centralization and decentralization along a few major periods: 1) mainframes and batch transaction processing (e.g. financial systems), fully centralized IT, end-users receiving outputs; 2) mainframes and online transaction processing, IT still centralized but end-users interacting with the system (e.g. ATMs, online reservation systems); 3) PCs, end-user computing and internal business decentralization; 4) Web 1.0, mass decentralization and full access to e-mail, home banking, online shopping, social interaction, etc.; 5) Web 1.0 plus outsourcing, where the front end of the business moves to the web, with non-competitive transaction processing systems and support being commoditized and located anywhere; and 6) Web 2.0 plus cloud computing, with virtualized organizations using web 2.0 tools, net PCs, mobile technology and cloud computing services.

After this contextual overview, we examine the definitions of cloud computing and its major components: hardware (INTEL, IBM chips to support virtualization); services (Google, Amazon.com); applications (SaaS, Software-as-a-Service) and virtualization (VMware), or

their combination (Citrix). We then explore the managerial implications of cloud computing and conclude by arguing that cloud computing represents a major IT change, transforming the way IT professionals work, and also a potential managerial revolution, with a fundamental change in how managers conceptualize and conduct business.

INFORMATION TECHNOLOGY PHASES

In order to better understand how cloud computing fits in the pendulum of centralization and decentralization of Information Technology, we should briefly examine some major periods in the last four decades of evolution of IT in organizations.

The first period was the 1970's era of mainframes and batch transaction processing. IT was fully centralized, and transactions related to payroll, financial statements, billing, accounting systems and others were processed in batches on the mainframe, offline, with the end-users simply receiving the outputs (printouts, reports).

The second period started in the 1980s, as transaction processing moved to online processing (e.g., credit cards, ATMs, online reservation systems). Point-of-Service (POS) terminals became ubiquitous and EDI use (electronic data interchange) became widespread. During this period, transactions were still centralized and still performed on the mainframe, with the difference that the submission interface was now online and users could interact directly with the system by performing queries and getting reports.

The third period happened in the 1980s and 1990s, with the PC Revolution, the explosion of end-user computing and internal business decentralization. Users stored data and ran applications on their own desktops or on their company's network. Initially they did all their computing at work, but eventually home computing came along, and users were able to use their home PCs to run simple applications like word processing and spreadsheets and perform small transactions.

By the mid 1990's, businesses started to grasp the IT potential of the World Wide Web, but structural and technical hurdles still remained before they could fully utilize this potential. In the late 1990's, however, capital markets caught the IT fever. Venture capitalists became eager to spend on IT, even when the long-term path to profitability was not clear. This led to the burst of the speculative bubble in the early 2000s, starting a downward spiral in IT that lasted until about 2003. The Web 1.0 represented the fourth period in IT evolution, bringing mass decentralization and giving everyone with Internet access the ability to conduct personal and work activities online: e-mail, home banking, online shopping, social interaction, etc.

The fifth period was the combination of Web 1.0 and outsourcing. The front end of business moved to the web, while the back end was outsourced- i.e., nonstrategic transaction processing systems, web support, anything that could be commodifized and done elsewhere in the world at a lower cost, started being seen as "services" that could be bought from outside providers who could be anywhere (on-shore in the US, near-shore in places like Mexico, Canada and Central America, or offshore in countries like China, India and Brazil). Outsourcing of IT tasks and computer services that could be clearly defined and were not part of the strategic core business allowed organizations to transform high IT fixed costs into lower outsourced variable costs. On the other hand, outsourcing alliances brought their own problems, because they often involved long-term contracts in an industry with a very high rate of technology change, and where few qualified players (e.g. EDS, CSC, IBM, ADP) had enough stature to handle the large outsourcing partnerships of the 1990's and 2000's. Managing IT outsourcing partnerships or alliances became a very complex process. CIOs found themselves in a position that was even more demanding than before: not only were they still responsible for the IT functions that remained in-house, but they also became responsible for planning, controlling and supervising the delivery of the outsourced IT services, while no longer having direct authority over these resources. Measuring performance now involved measuring both success and failure, and also determining the responsibility for failure in an environment where finger-pointing was common between clients and outsourcers.

The sixth and most recent period is the combination of Web 2.0 plus cloud computing. This means going beyond outsourcing, because both the front end and some of the back end of business can be outsourced. Instead of virtual organizations, we have virtualized organizations, with teams located anywhere in the world collaborating through the use of web 2.0 tools, net PCs, mobile technology and cloud computing services.

DEFINING CLOUD COMPUTING

The origin of the term "cloud computing" is nebulous. The term only gained traction around 2006 or 2007, but we found references dating back from much earlier. For example, a 1997 MIT paper (Gillett & Kapor [6]) showed a figure about the Internet's confederation approach, with the drawing of a cloud (labeled "cloud" of intermediate networks), to which originating and receiving networks were connected through routers. It became commonplace in the industry to use diagrams where the drawing of a cloud represented an Internetbased network where someone else is taking care of X, and I want to use X as a service without having to host it on my own server. While this kind of drawing was commonplace, the term "cloud computing" took a while to become widespread.

For many years, cloud computing remained "a collection of related concepts that people recognized, but didn't really have a good descriptor for, a definition in search of a term, you could say" (Willis [13]). One of the most cited examples of the first high profile public use of the term was in August 2006, when Google CEO Eric Schmidt used "cloud computing" at a search engine conference to describe what they were doing in terms of Software as Service (SaaS), which led the term to pick up "the PaaS/IaaS connotations associated with the Google way of managing data centers and infrastructure" (Willis [13]). Weeks later, Amazon used the word "cloud" when it launched its EC2 "elastic computing cloud" services, and the term entered the mainstream.

Currently there are multiple definitions of cloud computing, varying from very broad to very narrow, and emphasizing the perspective of different stakeholders. When the 2009 World Economic Forum started a research project about cloud computing, they had to grapple with the complexity and controversy of existing definitions, and ended up by adopting a broad definition of cloud computing that included "all kinds of remote services, from Software-as-a-Service to virtual machines." (Oram [10]).

Broad definitions often focus on the user perspective, in terms of what cloud computing allows individuals and organizations to do:

"In short, the cloud is the Real Internet, or what the Internet was really meant to be in the first place: an endless computer made up of networks of networks of computers. Even shorter: the Cloud is the Computer." Fingar [4]

"Cloud computing is the distributed virtualization of an organization's computing infrastructure." (Cagle [2]

"Applications and files are hosted on a "cloud" consisting of thousands of computers and servers, all linked together and accessible via the Internet. With cloud computing, everything you do is now web based instead of being desktop based. You can access all your programs and documents from any computer that's connected to the Internet." Miller [8]

"Gartner defines cloud computing (hereafter referred to as "cloud") as a style of computing where massively scalable IT-related functions and information are provided as a service across the Internet, potentially to multiple external customers, where the consumers of the services need only care about what the service does for them, not how it is implemented. Cloud is not an architecture, a platform, a tool, an infrastructure, a Web site or a vendor. It is a style of computing. Many architectures can be used to support its implementation and use. For example, it is possible to use cloud in private enterprises to build private clouds, but there is only one public cloud based on the Internet." Gartner Research [5]

By comparison, narrower definitions tend to focus on the technical aspects of the cloud:

"Cloud computing is grid computing, the use of a distributed network of servers, each working in parallel, to accomplish a specific task. As an acquaintance of mine put it, if it isn't using MapReduce, it probably isn't a cloud." Cagle [2]

A comprehensive review conducted in 2009 by the University of California Berkeley RAD Lab (Reliable Adaptive Distributed Systems Laboratory) yielded a definition that has been gaining broad popularity:

> "Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services. The services themselves have long been referred to as Software as a Service (SaaS). The datacenter hardware and software is what we will call a Cloud." Armbrust et al. [1]

James Urquhart [11] points out that there has been already too much energy spent on competing definitions and taxonomies of cloud computing, and that for the most part people basically understand what it is, and are ready to "change the conversation" to concentrate on its business aspects. Paraphrasing Supreme Court Justice Potter Stewart in the famous 1964 court case on obscenity and the First Amendment, Urquhart concludes that "the market seems to have come to the conclusion that cloud computing has a lot in common with obscenity - you may not be able to define it, but you'll know it when you see it."

Now that the U.S. Federal government has decided to use cloud computing, we may indeed be witnessing the "change in the conversation," that has been urged by Urquhart. After extensive consultations with IT industry experts and other stakeholders, the Information Technology Laboratory of the National Institute of Standards and Technology (NIST), a non-regulatory agency of the Commerce Department, issued what is expected to become the "de facto standard definition" (Urquhart [12]) of cloud computing. The NIST formal multi-part definition starts with two precautionary notes, and defines cloud computing by delineating five essential characteristics, three service models and four deployment models, as follows:

> "Note 1: Cloud computing is still an evolving paradigm. Its definitions, use cases, underlying technologies, issues, risks, and benefits will be refined in a spirited debate by the public and private sectors. These definitions, attributes, and characteristics will evolve and change over time.

> Note 2: The cloud computing industry represents a large ecosystem of many models, vendors, and market niches. This definition attempts to encompass all of the various cloud approaches.

Definition of Cloud Computing:

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

Essential Characteristics:

- On-demand self-service. A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.
- Broad network access. Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).

- Resource pooling. The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.
- Rapid elasticity. Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.
- Measured Service. Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service.

Service Models:

- Cloud Software as a Service (SaaS). The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating svstems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.
- Cloud Platform as a Service (PaaS). The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications

created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

Cloud Infrastructure as a Service (IaaS). The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

Deployment Models:

- Private cloud. The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.
- Community cloud. The cloud infrastructure is shared by several organizations and

supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

- Public cloud. The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.
- Hybrid cloud. The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).
- Note: Cloud software takes full advantage of the cloud paradigm by being service oriented with a focus on statelessness, low coupling, modularity, and semantic interoperability." Mell & Grance [7]

Figure 1 shows the service models with another level including business process as a service, beyond the original service model definition.

Figure 2 shows a representation of a future scenario of the deployment models for Corporate and Federal agencies, beyond the original definition.

Business Process (BPaaS)	SaaS plus business process customization by provider
Software as a Service (SaaS)	Applications to be customized and used in provider site
Platform as a Service (PaaS)	Tools and languages for application development with DB and integration support
Infrastructure as a Service (IaaS)	Servers, network and storage hardware facilities

Figure 1: Cloud Computing Service Models

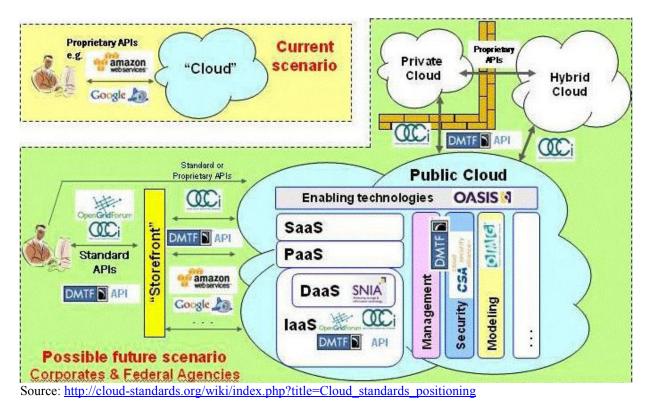


Figure 2: Deployment Models for Corporate and Federal Agencies

WORKING IN THE CLOUD

The major components and industry players in cloud computing include: hardware (INTEL, IBM chips to support virtualization); software (VMware, Microsoft, etc.); services (Google, Amazon.com); applications (Software-as-a-Service, or SaaS); virtualization (VMware) or their combination (e.g., Citrix, a classic case of virtualization that merges SaaS and virtualization).

A service provider (such as Google or Amazon) provides hardware and virtualization software, and sometimes also applications. Instead of us, as users, hosting our own servers, their computers run virtual machines where our server can reside. The service provider's machine has a certain type of software that does virtualization (VM ware), so that a single machine in the provider can run many virtual machines, each with own operating system, hard drives, and application software. If virtualization were done only through software, it would be very slow, so chip and CPU manufacturers like INTEL and AMD embedded in their chip set hardware that allows virtualization support to create virtual machines. To avoid the security problems of virtual machines, manufacturers such as INTEL make the chip CPU with a special code so that when we, as users, log into the virtual machine in the provider's physical machine, we can be sure that the host machine is indeed ours, and not accessible to others without our permission.

It should be noted that cloud computing can also be done in-house: instead of an organization having PCs in people's desks, there may be just netbooks connected to the cloud, in recognition of the fact that most of the time the computing capacity of isolated PCs is just being wasted.

MANAGERIAL IMPLICATIONS

In the first period of IT evolution we examined here (mainframes and batch transaction processing, fully centralized IT, end-users receiving outputs), computers existed in a "secret world" separate from users, who were not familiar with them as physical objects, nor with their operations and jargon. During the second period (mainframes and online transaction processing, IT still centralized), computers became more of a visible entity, as end-users started interacting with them through interfaces such as ATMs and online reservation systems; the nature of doing business was transformed by IT, but that transformation did not reach management, who could still consider that technology was someone else's problem. This management insulation changed in the third period (PCs, end-user computing, EDI), with internal business decentralization and management's realization that they were now responsible for managing not only their own organization, but a network of interorganizational relationships and partnerships with customers and suppliers.

During the fourth period (Web 1.0, mass decentralization and full access to e-mail, home banking, online shopping, social interaction, etc.), the web dramatically reduced the costs of EDI-like partnerships, making it possible for businesses of all sizes to have a broad web presence. Many of them, however, still maintained "walls" between their online and brick-andmortar operations, and had to learn some hard lessons as they climbed the learning curve of thinking as a seamless organization (for example, Toys-R-us had such an unexpected success when first selling toys online for Christmas that they run out of inventory and were late in their Christmas deliveries, because their online and brickand-mortar inventories were run separately).

In the fifth period (Web 1.0 plus outsourcing), the front end of the business moved to the web, with the commoditization and outsourcing of non-competitive transaction processing systems and support. As consumers gained the ability to use social electronic media in every aspect of their lives, e-commerce gained an unprecedented impetus, and both small businesses and large companies became denizens of the web.

The sixth period (Web 2.0 plus cloud computing) is opening an era of fundamental managerial changes of business organizations, with virtualized organizations using web 2.0 tools, net PCs, mobile technology and cloud computing services.

Miller [8] and Carr [3] compare the advent of cloud computing to last century's electricity revolution. Instead of having to produce electricity from their own individual generators, businesses became able to buy power from electrical utilities, which both cut costs and improved reliability. By the same token, cloud computing liberates businesses from having to generate and manage their own computing power, frees them from the mainframe and desktop-centric frameworks of the past and opens a future where they can expect universal, 24/7 access to computing resources that someone else is providing and managing in the cloud.

In such a world, virtualized organizations rely on teams that use Web 2.0 and the cloud to collaborate anywhere, anytime. This is not just IT change, but a potential management revolution. Fingar [4] proposes that

with the cloud "the world shifts from using Information Technology (IT) for transaction and information management to a far more organic Business Technology (BT) for collaboration and interaction management." The question becomes: are managers ready to lead and manage on the web, using web 2.0 tools and cloud computing services to generate innovation? Elements of the business infrastructure that have always been taken for granted, such as time and space requirements, are suddenly up for grabs. If an organization's data, files, programs, applications are all in the cloud, there is no longer the need for many local machines and hard drives, and massive decentralization becomes possible, as employees just need a netbook, digital tablet, or even a smart phone to store, retrieve and work collaboratively wherever they are – at work, at home, or on the road.

Increasingly, these employees are "digital natives," who grew up using laptops and smart phones, who prefer instant messaging to e-mail, and whose first reaction when they don't know something is to tap into a network of those who might know it, whether or not those people are friends or strangers, inside or outside the organization (Ommeren et al. [9]). Leading and managing these digital natives will require older managers to become "naturalized" citizens of this sixth period in the evolution of IT in organizations, who can themselves "live," personally and professionally, in what might seem to them the foreign land of Web 2.0 and cloud computing.

Cloud computing changes the way IT professionals will work, and the kinds of jobs they will have. But it also brings a fundamental change in how managers think about business, coordinate tasks and people. Fingar [4] points out that "in a process-managed enterprise, command-and control leadership gives way to connect-and-collaborate, where every member of a business team is a leader. It's about acting on opportunities, and letting others lead the leader when they know best about stuff being done (...). Although the Cloud enables radical change, the *culture* of the firm will determine the outcome. Permission, risk tolerance, cultivating lots of small bets – these are some of the earmarks of a Cloud-oriented business culture".

As "immigrants" to the brave new world of Web 2.0 and the cloud, managers' biggest challenge is to learn how the "digital natives" think, learn and act, so that they can let them – and their organizations - soar.

REFERENCES

 Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R, Konwinski, A., Lee, F., Patterson, D., Rabkin, A., Stoica, I. & Zaharia, M. "Above the clouds: A Berkeley view of cloud computing." University of California Berkeley, Reliable Adaptive Distributed Systems Laboratory, Technical Report No. UCB/EECS-2009-28, (2009). Accessed January 20, 2011 [available at http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/ EECS-2009-28.html]

- [2] Cagle, K. "But what exactly 'is' cloud computing?" O'Reilly Broadcast, 2008. Accessed January 20, 2011 [available at <u>http://broadcast.oreilly.com/2008/12/but-what-exactly-is-cloud-comp.html</u>]
- [3] Carr, N. *The big switch: rewiring the world, from Edison to Google*. W.W. Norton & Co., New York, NY, 2008.
- [4] Fingar, P. "Cloud computing set to unleash a perfect storm in business." *Cordial Cloudburst*, 2009. Accessed January 20, 2011 [available at <u>http://www.cordys.com/ufc/file2/cordyscms_sites/d</u> <u>ownload/6f5f4d1cfe8be9d78d972fa808d8702c/pu/c</u> ordial fingar.pdf]
- [6] Gillett, S. E. & Kapor, M. "The self-governing Internet: Coordination by Design." In: *Coordination of the Internet*, Brian Kahin and James Keller (Eds.). MIT Press, Cambridge, MA, 2008.
- [7] Mell, P. & Grance, T. *The NIST Definition of Cloud Computing*. U.S. Department of Commerce, National Institute of Standards and Technology, Information Technology Laboratory. Version 15, 10-7-09, 2009. Accessed January 20, 2011 [available at <u>http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc</u>]
- [8] Miller, M. Cloud computing: Web-based applications that change the way you work and collaborate online. Que Publishing, Indianapolis, IN, 2008.
- [9] Ommeren, E. V., Duivestein, S., deVadoss, J, Reijnen, C. & Gunvaldson, E. *Collaboration in the Cloud*. Microsoft and Sogeti, Bariet, Ruinen, the Netherlands, 2009.
- [10] Oram, A. "Cloud computing perspectives and questions at the World Economic Forum," *WikiContent*, 2009. Accessed January 20, 2011 [available at

http://commons.oreilly.com/wiki/index.php?title=C loud_computing_perspectives_and_questions&prin table=yes]

- Urquhart, J. "The cloud conversation is changing." *CNET News: The wisdom of clouds*, June 6, 2009. Accessed January 20, 2011 [available at <u>http://news.cnet.com/8301-19413_3-10258721-</u> <u>240.html</u>]
- Urquhart, J. "Are the Feds the first to a common cloud definition? *CNET News: The wisdom of clouds*, May 10, 2009. Accessed January 20, 2011 [available at <u>http://news.cnet.com/8301-19413 3-10237274-240.html</u>]
- [13] Willis, J. "Did Google's Eric Schmidt coin 'cloud computing'?" Virtualization Journal, 2009. Accessed January 20, 2011[available at <u>http://virtualization.sys-con.com/node/795054</u> and <u>http://www.johnmwillis.com/cloud-computing/</u> who-coined-the-phrase-cloud-computing/]

ABOUT THE AUTHORS

Al Bento is Professor of Information Systems, University of Baltimore. He was previously on the faculty of Boston University School of Management, California State University and Bentley University. He has served as Department Chair and Research Center Director in these institutions. Prior to his academic career, he worked for nine years at IBM World Trade as Systems Engineer, Budget Manager, and Education and Scientific Affairs Manager. He has served the profession in a variety of capacities, including Chair of the Association for Information Systems (AIS) Special Interest Group on Security (SIGSEC). He holds a B.S.B.A, an M.S. in Computer Science and Systems Engineering, and a Ph.D. in Management, Computer Information Systems from the University of California, Los Angeles (1980).

Regina Bento is Professor of Management at the Merrick School of Business, University of Baltimore. After graduate studies at UFRJ, Federal University of Rio de Janeiro, she pursued doctoral studies at Harvard and MIT (Ph.D. MIT, 1990). She taught at UFJR (1980-1982) and UC Riverside (1988-1991) before joining UB in 1991. She was a Visiting Professor at MIT (1998, 2006), and Associate Director at the Christensen Center, Harvard Business School (2006 to 2009). She has served in leadership positions at the Academy of Management, and on numerous editorial boards. Regina has received the USM Regents Award, the highest honor in the University System of Maryland.