The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research

Introduction

The miniaturization of hardware, increasingly powerful microprocessors, inexpensive and reliable memory, broadband communication, and efﬁcient power management have made it possible to digitize key functions and capabilities of industrial-age products including cars, phones, televisions, cameras, and even books (Yoo 2010). With embedded digital capability, such products offer novel functions and remarkably improved price/performance ratios that transform their design, production, distribution, and use. The phenomenal success of Apple’s iPhone and Amazon’s Kindle exempliﬁes how the digitization of well-established products such as books sparks profound changes in the industrial structure and competitive landscape, blurring industry boundaries and creating new threats and opportunities. In the e-book case, ﬁrms from the computer industry, consumer electronics, Internet search, online retailing, book retailing, telecommunications, and publishing form dynamic and overlapping alliances that are being mingled together into a complex ecosystem. In this ecosystem, ﬁrms are busily developing new strategies that cater for the emerging market dynamics by competing head-to-head on some fronts (e.g., both Apple and Amazon sell hardware) and collaborating on others (e.g., Amazon offers reader applications for Apple’s iPad). The digitization of the book is fundamentally reshaping the structure that has underpinned book publishing for 200 years by bringing together ﬁrms from previously unrelated industries, ultimately changing the very idea of a book.

Over the last decade, information systems (IS) scholars have successfully examined the impacts of digital technology on ﬁrms’ strategies, structures, and processes (Sambamurthy et al. 2003, Sambamurthy and Zmud 2000). Similar advances have been made to understand the role of information technology (IT) in creating business value and building sustainable competitive advantage (Kohli and Grover 2008, Nevo and Wade 2010). However, digital technology’s transformative impact on industrial-age products has remained surprisingly unnoticed in the IS literature. In fact, the IS literature rarely considers how product architectures—the arrangement of functional elements, the mapping from functional elements to physical components, and the speciﬁcation of interfaces among components (Ulrich 1995, p. 420)—affect a ﬁrm’s strategic choices and related IT deployments. Neither has the literature considered the emergence of new organizing logics—i.e., the “managerial rationale for designing and evolving speciﬁc organizational arrangements in response to an enterprise’s environmental and strategic imperatives” (Sambamurthy and Zmud 2000, p. 107)—spurred by changes in product architecture because of digital technology. This is unfortunate because changes in product architecture and organizing logic reshape the landscape of IS strategy and use in ﬁrms.

In this essay, we propose that digital technology instigates a new type of product architecture: the layered modular architecture. We conceive layered modular architecture as a hybrid of the modular architecture of a physical product and the layered architecture of digital technology. The modular architecture provides a scheme by which a physical product is decomposed into loosely coupled components, is attributed functionality, and is then interconnected through pre-speciﬁed interfaces (Baldwin and Clark 2000, Ulrich 1995). The layered architecture of digital technology (Adomavicius et al. 2008, Gao and Iyer 2006) is embedded into physical products, enhancing product functionality with software-based capabilities. Similar to modularity’s impact on industrial organization (Baldwin and Clark 2000, Langlois 2003), we argue that the emergence of layered modular architecture generates profound changes in a ﬁrm’s organizing logic and innovation. To this end, we (1) develop a conceptual framework to characterize the organizing logic of digital innovation based on layered modular architecture, and (2) formulate an IS research agenda to study the new logic and its effects on digital strategy and corporate IT infrastructures.

Digital Innovation

Deﬁning Digital Innovation

Following Schumpeter (1934), we deﬁne digital innovation as the carrying out of new combinations of digital and physical components to produce novel products. Our use of the term digital innovation thus implies a focus on product innovation, distinguishing it from extant IT innovation research that has been primarily occupied with process innovation (Swanson 1994). A necessary but insufﬁcient condition for digital innovation is that the new combination relies on digitization, i.e., the encoding of analogue information into digital format. Digitization makes physical products programmable, addressable, sensible, communicable, memorable, traceable, and associable (Yoo 2010). Digital innovation furthermore requires a ﬁrm to revisit its organizing logic and its use of corporate IT infrastructures.

Consider the e-book example: Digitization has created a necessary condition for digital innovation among a range of ﬁrms capable of engaging in digital publishing. The previously non-digital product— the book—now embeds digital capabilities such as communication, memory, programmability, traceability (e.g., Amazon can track how long readers look at pages and readers can ﬁnd out who else underlined particular sentences), and so on. Despite the short history of the e-book, there are already signs of changes in the organizing logic of publishing whereby publishers’ tight control over the content creation, production, and distribution is deteriorating. In the early stages of the e-book evolution, Kindle replaced an old physical artefact with a new one with similar (although digitally enabled) form factors. Kindle’s main attractions were the radically reduced marginal production and distribution costs and its ability to hold thousands of books in a single unit. With the introduction of iPad some 18 months later, however, e-books challenge the vertically integrated model of publishing. The e-book is now fully disintegrated into distinct layers of devices, networks, services, and contents—a fate already experienced by the digital camera and mobile phone, and likely to be repeated with television with products such as Google TV and Apple’s iTV. Following the disintegration of the vertical model, new conceptions of a book are likely to sprout as other digital components such as interactive multimedia, GPS, social media applications, and accelerometers are being integrated into e-books.

Key Characteristics of Digital Innovation

In order to understand the nature of digital innovation, one must consider how digital technology differs from earlier technologies. Here, we note three unique characteristics: (1) the re-programmability, (2) the homogenization of data, and (3) the self-referential nature of digital technology.

First, based on the von Neumann architecture, a digital device consists of a processing unit that executes digitally encoded instructions and a storage unit that holds both instructions and the data being manipulated in the same format and in the same locations (Langlois 2007). As long as users agree on the meaning of the digital data and have the wits to come up with new instructions to manipulate the data, the architecture offers ﬂexibility in the way data is manipulated. Thus, unlike analogue technology, a digital device is reprogrammable, enabling separation of the semiotic functional logic of the device from the physical embodiment that executes it. The re-programmability allows a digital device to perform a wide array of functions (such as calculating distances, word processing, video editing, and Web browsing).

Second, an analogue signal maps changes in a continuously varying quantity on changes in another continuously changing quantity. As such, analogue data implies a tight coupling between data (e.g., texts and pictures) and special purpose devices for storing, transmitting, processing, and displaying the data (e.g., book and camera). In contrast, a digital representation maps any analogue signal into a set of binary numbers, i.e., bits (a contraction of binary digits). This leads to a homogenization of all data accessible by digital devices. Any digital contents (audio, video, text, and image) can be stored, transmitted, processed, and displayed using the same digital devices and networks. Furthermore, unlike analogue data, digital data originate from heterogeneous sources and can be combined easily with other digital data to deliver diverse services, which dissolves product and industry boundaries. Thus, the homogenization of data along with the emergence of new media separates the content from the medium.

Finally, self-reference means that digital innovation requires the use of digital technology (e.g., computers). Therefore, the diffusion of digital innovation creates positive network externalities that further accelerate the creation and availability of digital devices, networks, services, and contents (Benkler 2006, Hanseth and Lyytinen 2010). This, in turn, fosters further digital innovation through a virtuous cycle of lowered entry barriers, decreased learning costs, and accelerated diffusion rates. The drastic improvements in the price/performance of computers and the emergence of the Internet have made the digital tools necessary for innovation more affordable to a broad spectrum of previously excluded economic and innovative activity. Digital technology, therefore, has democratized innovation and almost anyone can now participate.

The Layered Architecture of Digital Technology

The characteristics of digital technology pave the way for layered architecture (Adomavicius et al. 2008, Gao and Iyer 2006) and this is perhaps best exempliﬁed by the Internet. The layers manifest two critical separations: (1) that between device and service because of re-programmability and (2) that between network and contents because of the homogenization of data.

As illustrated in Figure 1, layered architecture consists of four layers: devices, networks, services, and contents (Benkler 2006, Farrell and Weiser 2003). The device layer can be further divided into a physical machinery layer (e.g., computer hardware) and a logical capability layer (e.g., operating system). The logical capability layer provides control and maintenance of the physical machine and connects the physical machine to other layers. The network layer is similarly divided into a physical transport layer (including cables, radio spectrum, transmitters, and so on) and a logical transmission layer (including network standards such as TCP/IP or peer-to-peer protocols). The service layer deals with application functionality that directly serves users as they create, manipulate, store, and consume contents. Finally, the contents layer includes data such as texts, sounds, images, and videos that are stored and shared. The contents layer also provides metadata and directory information about the content’s origin, ownership, copyright, encoding methods, content tags, geo-time stamps, and so on.

The four layers represent different design hierarchies (Clark 1985), and the individual design decisions for components in each layer can be made with minimum consideration of other layers. Therefore, designers can pursue combinatorial innovation by gluing components from different layers using a set of protocols and standards to create alternative digital products (Gao and Iyer 2006). Combined with the rapid diffusion of personal computers and the Internet, the layered nature of digital technology has brought unprecedented levels of generativity (Tuomi 2002, Zittrain 2006).

Though layered architecture has been discussed in IS literature (Adomavicius et al. 2008, Gao and Iyer 2006), little attention has been paid to its implications for product innovation. The digitization of physical products challenges some of the fundamental assumptions about product architecture and organizing logics. Next, we will discuss how it introduces a new type of product architecture.

Layered Modular Architecture

Modular Architecture

Two architectures have dominated physical product design: integral and modular. An integral architecture is characterized by a complex and overlapping mapping between functional elements and physical components, where the interfaces between components are not standardized and are tightly coupled (Ulrich 1995). As a result, changes in one part of a product typically affect the rest of the product, often unpredictably. The tight coupling among components in an integral architecture renders high performance and quality, which is important for certain products such as sports cars and high-end electronics.

Conversely, a modular architecture is characterized by its standardized interfaces between components. Modularity is a general characteristic of a complex system and refers to the degree to which a product can be decomposed into components that can be recombined (Schilling 2000). Rooted in Simon’s (1996) design theory, modular architecture offers an effective way to reduce complexity and to increase ﬂexibility in design by decomposing a product into loosely coupled components interconnected through pre-speciﬁed interfaces (Baldwin and Clark 2000). Although just “nearly decomposable” in practice (Simon 2002), an ideal modular architecture implements one-to-one mapping between functional elements and physical modules (Ulrich 1995).

Shifts in product architecture cause shifts in the organizing logic of a ﬁrm. With an integral product architecture, the dominant organizing logic is the vertically integrated hierarchy, wherein a single ﬁrm carries out the majority of innovation required to compete. Here, components are often co-specialized with each other (Langlois 2003, Teece 1993). The key sources of value creation are economies of scale and scope, which emanate from overwhelming endowments to physical resources (Barney 1996). With an integral architecture, dominant approaches to competitive strategy are product positioning (Porter 1980), which distinguishes market scope and strategic strength as key parameters for determining the appropriate strategy. In contrast, a modular architecture leads to vertical disintegration of a ﬁrm’s design and production functions, as seen in the change of the industrial organization of the computer (Baldwin and Clark 2000, Langlois 2007), software (Chandler and Cortada 2000), and telecommunication industries (Tuomi 2002). Leveraging radically reduced communication and coordination costs enabled by IT (Malone 2004), ﬁrms such as Cisco, Dell, and Nokia have heavily invested in corporate IT infrastructures in order to realize net-enabled value networks (Sambamurthy and Zmud 2000, Wheeler 2003). This enabled them to distribute design and production activities among a network of ﬁrms (Nohria and Eccles 1992). The key source of value creation is the agility that ﬂows from the ability to rapidly recombine components of a modular product architecture positioned within a single design hierarchy without sacriﬁcing cost or quality (Sambamurthy et al. 2003).

Layered Modular Architecture

As ﬁrms increasingly embed digital components into physical products, the layered modular architecture emerges. The layered modular architecture is a hybrid between a modular architecture and a layered architecture, where the degree by which the layered architecture adds the generativity to the modular architecture forms a continuum. At one end, we have the traditional modular architecture based on a ﬁxed product boundary. The modular design of such a product is initiated by decomposing the product into components following a functional design hierarchy (Clark 1985, Baldwin and Clark 2000). Therefore, the relationships between the product and its components are nested and ﬁxed. Given the nested nature of relationships and the ﬁxed product boundary, aggregating all components will make up the whole product. In addition, in a modular architecture, the design of a component is driven by the functional requirements created within the context of a given product. That is, components in a modular architecture are product speciﬁc. Furthermore, components are designed and produced by specialized ﬁrms that all share product-speciﬁc knowledge. The primary goal of modularity is to reduce complexity and to increase ﬂexibility (Schilling 2000, Simon 1996). The ﬂexibility is accomplished through substitutions of components within a single design hierarchy. For example, a single lens reﬂex (SLR) camera can be ﬁtted with multiple lenses using a standardized mounting interface, which increases the camera’s ﬂexibility. Thus, the ﬂexibility of a modular architecture comes from the differences in degree.

At the other end, we have the full-blown layered modular architecture that does not have a ﬁxed boundary at the product level. The design of a component thus requires little product-speciﬁc knowledge. That is, components in a layered modular architecture are product agnostic. Google Maps, for example, consists of a bundle of contents (i.e., maps) and service (e.g., search, browse, trafﬁc, and navigation) layers with different sets of interfaces (i.e., application programming interfaces). Though Google Maps can be used as a standalone product, it can simultaneously be used in a variety of different ways, bundled with a host of heterogeneous devices such as desktop computers, mobile phones, televisions, cars, navigation systems, or digital cameras. In this regard, a component design in a layered modular architecture is not derived from a single design hierarchy of a given product. Instead, a product is inductively enacted by orchestrating an ensemble of components from a set of heterogeneous layers, each of which belongs to a different design hierarchy (Clark 1985). Therefore, the designers of components in a layered modular architecture cannot fully know how the components will be used. That is, Google’s designers cannot fully anticipate all the possible ways that Google Maps as a component will be used. As such, a layered modular architecture offers generativity, i.e., “a technology’s overall capacity to produce unprompted change driven by large, varied, and uncoordinated audiences” (Zittrain 2006, p. 1980). Generativity in a layered modular architecture is accomplished through loose couplings across layers whereby innovations can spring up independently at any layer, leading to cascading effects on other layers (Adomavicius et al. 2008, Boland et al. 2007). Whereas components in a modular product fall under a single design hierarchy, components in a layered modular architecture participate in multiple heterogeneous design hierarchies. Unlike the ﬂexibility of a modular product that produces differences in degree, the generativity of a layered modular product produces differences in kind. For example, in utilizing available hardware resources, a digital camera with a layered modular architecture can be used not only as a camera but also as a video player, photo editor, Internet client, and in many other ways. Therefore, a layered modular product remains ﬂuid and is open to new meanings. Unlike the purely layered architecture (Gao and Iyer 2006), however, the generativity of a digitized product with a layered modular architecture is constrained by characteristics of the physical components of the product (e.g., form factors and availability of certain physical components).

The modular architecture and the layered modular architecture form the two end points of a continuum as ﬁrms embed digital components (see Figure 2) into their products. Traditional industrial-age, single purpose products manifest one end of the spectrum while conventional digital products with general computer hardware form another end. Many digitized products will fall somewhere in the middle.

The Organizing Logic of Layered Modular Architecture

With a layered modular architecture, a digitized product can be simultaneously a product and a platform. For instance, an iPad can be used as a complete product out of the box. Yet, as a platform, it enables other ﬁrms to invent novel components such as new applications and peripheral hardware accessories with which its basic functionality can be expanded. Therefore, ﬁrms operating in a competitive landscape shaped by layered modular architectures invest in digital product platforms that cater for multisided markets and help build vibrant ecosystems (Eisenman et al. 2006). A digital product platform typically encompasses a particular range of layers (e.g., content and service layers) that can function as a new product, but simultaneously enable others to innovate upon (Gawer and Cusumano 2008) using ﬁrm-controlled platform resources (e.g., SDKs and APIs2). For example, as most subsystems of an automobile are becoming digitized and connected through vehicle-based software architectures, an automobile has become a computing platform on which other ﬁrms outside the automotive industry can develop and integrate new devices, networks, services, and content (Henfridsson and Lindgren 2010).

A digitized product with a layered modular architecture can serve as a platform courting for its own installed base at one layer and serve as a component at another layer. Because of the dynamic nature of the layered modular architecture, the same ﬁrms can compete on one layer and peacefully coexist on other layers. For example, Apple’s iPad and Amazon’s Kindle directly compete at the device layer. The two ﬁrms also compete at the content layer with iBook and Kindle stores. At the same time, Amazon offers an application for iPad and is thus a component provider at the service layer of the iPad. Similarly, Apple’s iPhone (device layer), along with other mobile devices, has been an important component for Google’s mobile search platform (service layer). At the same time, Google Maps (service and content layers) is an important component of the iPhone platform. As Apple introduced its own mobile search and advertisement systems, however, Apple and Google began to compete directly on the service layer. Similarly, when Google introduced its own Android-based mobile phone, Apple and Google began to compete at the device layer.

Within a layered modular architecture, a ﬁrm seeks to attract heterogeneous actors to design and produce novel components on layers outside of its digital product platform. The generativity of a layered modular architecture thus comes from a ﬁrm’s ability to design a product platform that can attract a large number of heterogeneous and unexpected components that belong to different design hierarchies. The greater the heterogeneity, the more generative the platform becomes. Although it is theoretically possible to pursue such generativity within the closed boundary of a single ﬁrm or its existing supplier network, a ﬁrm’s ability to do so on a practical basis is limited by its economic, structural, cognitive, and institutional constraints. Therefore, even though the layered modular architecture may be ripe with generative potential, this potential is only fully realized when it is paired with a new organizing logic that involves heterogeneous actors, many of whom pursue their own innovation strategies. As a result, innovation within a layered modular architecture is distributed not only among ﬁrms of the same ilk but also across ﬁrms of different kinds. These ﬁrms’ innovation activities reciprocally and recursively inﬂuence each other, creating the image of “wakes of innovation” (Boland et al. 2007). Accordingly, we characterize the organizing logic for a layered modular architecture as doubly distributed. It is distributed because the primary source of value creation is the generativity that comes from the unbounded mix-and-match capability of heterogeneous resources across layers. It is doubly distributed because (a) the control over product components is distributed across multiple ﬁrms, and (b) the product knowledge is distributed across heterogeneous disciplines and communities. In this environment, an essential capability is the ability to design a digital product platform to inspire and mobilize a vibrant and doubly distributed network to maximize the generative potential of the layered modular architecture. In managing such a network, a ﬁrm needs to have the capability to create new meanings of its products and services (Verganti 2009) by constantly redeﬁning the product boundaries through active reshaping of the product ecology (Kusuoki and Aoshima 2010).