

THE ARCHITECTURE OF GENERATIVITY IN A DIGITAL ECOSYSTEM: A SYSTEMS BIOLOGY PERSPECTIVE

Research-in-Progress

SungYong Um

Fox School of Business,
Temple University
1810 North 13th Street, Philadelphia,
Pennsylvania 19122
sungyong.um@temple.edu

Bin Zhang

Fox School of Business,
Temple University
1810 North 13th Street, Philadelphia,
Pennsylvania 19122
bzhang@temple.edu

Rob Kulathinal

Department of Biology,
Temple University
1810 North 13th Street, Philadelphia, Pennsylvania 19122
robkulathinal@temple.edu

Sunil Wattal

Fox School of Business,
Temple University
1810 North 13th Street, Philadelphia,
Pennsylvania 19122
swattal@temple.edu

Youngjin Yoo

Fox School of Business,
Temple University
1810 North 13th Street, Philadelphia,
Pennsylvania 19122
youngjin.yoo@temple.edu

Abstract

Generativity of digital artifacts enables firms to leverage third party developers to build new functionalities for open digital ecosystems. This paper employs a systems biology approach to study how innovations from third party developers create generative momentum in such ecosystems. We propose a theoretical framework on how the structural pattern of the modules of the ecosystem and the relations among these modules impacts their generativity. We test this theoretical framework using data from WordPress.org, which is the world's largest blog service. Using source code from WordPress plug-ins, we extract the APIs of the plug-ins and construct a topological overlap plot of the ecosystem by identifying the common APIs among various plug-ins. We test the model by examining the concept of a gene co-expression network. In particular, we conceptualize the functionality of different digital artifacts as an expression of different interacting modules. This paper makes unique methodological and theoretical contribution by demonstrating the use of techniques from systems biology, in particular genetic networks, to study open innovation.

Keywords: modular innovation, generativity, a gene coexpression network, topologic overlap, module distribution

Introduction

Firms using digital technologies can no longer focus only on the design and control of their own platform as a way to offer innovative features. Instead, they try to leverage their open digital ecosystem by allowing third-party developers to build plug-ins that provide new features that were often not anticipated by the original platform owner (Sanchez 2000). The versatile application of various modules such as APIs helps third party developers to create a new digital artifact through recombination that leads to the blurring boundary of a digital ecosystem (Yoo et al. 2010). Traditionally, modules were designed and produced within a given product with a fixed boundary to allow flexibility without redesigning the system through substations of modules and subsystems (Garud and Kumaraswamy 1993, Sanchez and Mahoney 1996). However, emerging digital ecosystems with vibrant third-party developers community challenges this traditional concept of modularity based on a fixed system boundary (Zittrain 2006, Kallinikos and Mariategui 2011). These new digital ecosystems emphasize the generativity of third party developers who show explicit identification across the system boundary.

Generativity refers to the capacity of technology to produce unprompted changes driven by large, varied, and uncoordinated participants (Zittrain 2006). Generativity is derived from unique material characteristics of digital technology such as reprogrammability, data homogenization and self-reflexibility (Yoo et al. 2010). Such material characteristics of digital technology allow firms to create new products (re)combining as wide range of existing modules that are highly reprogrammable. The logic of generativity offers an insight on how organizations such as Apple and Google offer a large array of innovations through their platforms. This dynamic nature of innovations in digital ecosystems primarily derives from interactions among modules produced by third party developers rather than through coordinated integration by a central firm, which has architectural knowledge (Henderson and Clark 1992). However, little attention has been paid to explore the underlying structure of generativity and how innovations by some can instigate wakes of innovations through the entire digital ecosystem (Boland et al. 2007).

Information Science (IS) scholars have begun to study technology-enabled innovations in organizations (Swanson 1994, Samurthy 2003). However, most of these studies are confined within the context of firms that are innovating following the logic of modularity (Yoo et al. 2010). As such, existing theories on innovation in IS literature cannot effectively capture dynamic and relentless innovations that we see in contemporary digital ecosystems. Most of these studies focus on how information technology can improve the efficiency and effectiveness of an organization with a fixed boundary. What we need is a new theoretical approach that embraces the dynamic and complex nature of unbounded innovations that we observe in digital ecosystems.

In this paper, we draw on the models of generativity from the systems biology literature that emphasizes the interactions among modules (i.e., genes or proteins in biology) as a way of understanding the structure and dynamics of cells and organisms (Kitano 2002a). Particularly, the field of systems biology is concerned with how the interactions among functionally diverse set of genes produce coherent functions of a cell and organism. In this paper, we argue that models from systems biology can be effectively used to explain the structure and the dynamics of digital ecosystem using the interactions among diverse modules. Such an insight can help us gain systematic understanding of the generativity in digital ecosystems.

In systems biology, a gene co-expression network can show how various interactions among a diverse set of genes in a cell produces coherent functions. Using co-expression genetic networks, biologists cluster genes according their co-expression patterns (Ravasz et al. 2002, Stuart et al. 2003, Zhang and Horvath 2005, Ruan et al. 2010). Such a network is particularly effective in finding patterns of gene interactions when the number of genes is large. Similarly, we can identify patterns of (re)combinations among modules in a digital ecosystem as a way of understanding its structure and dynamics using a network model that is similar to a gene co-expression network. Specifically, using a gene co-expression network in digital ecosystems, we can ask:

- a) What are the underlying structural patterns of generativity in a digital ecosystem?
- b) How do new functions emerge in a digital ecosystem through (re)combinations of modules?

By answering these questions, we will be able to provide insights on how to design an ecosystem that is generative.

Literature Review

Modular Architecture

Simon (1962) proposed the notion of modularity as a way of dealing with a complex system through a hierarchical relationship between the system and its modules. One can use a modular architecture to represent a complex system divided into a set of subsystems that perform specific functions and their relationships (Ulrich 1995). As such, modularity offers simplicity in dealing with a complex system as one can focus on an overall architectural scheme that defines the boundary of the system, the modules that make up the system and interfaces among the modules, while leaving the detail design of the modules to others. This allows an effective division of labor among different actors during the design and production of a complex system (Sosa et al. 2004, Staudenmayer et al. 2005).

Not only modularity reduces the complexity in design and production of a product, it also increases its flexibility by allowing “mixing-and-matching” of various modules (Sanchez and Mahoney 1996). Such a “mixing-and-matching” strategy is possible as one can replace one module with another as long as they both confirm to the same standardized interface. For example, with person computers, a user can use any video board or a monitor as long as they follow the same standard and protocol. As such, firms can increase the product variations by leveraging modules increasingly “intermeshed” way (Sanchez 2004b).

Traditionally, scholars have emphasized the role of the firm who controls dominant design of a product through its architectural knowledge (Henderson and Clark 1990, Clark 1985). However, recently scholars started arguing that heterogeneous firms can autonomously coordinate (either explicitly or implicitly) their designs and productions of modules without the intervention of the architecture owner when the architecture is mature enough with a set of clearly articulated interfaces (Staudenmayer et al. 2005). In this sense, a product can be considered as a network of peer modules that share the same interfaces in order to function as a whole (Sosa et al. 2005). By leveraging such flexibility that come from modularity, a firm who owns the architecture can increase the productions variations and agility of organizations without incurring high transaction cost (Sosa et al. 2004, Sambamurthy et al. 2003).

While such a network perspective of modular innovations helps us understand production innovations, these models makes implicit assumptions that the boundary of system is fixed and clearly articulated. Therefore, the flexibility of a product and its variations take place within a confinement of a given architectural scheme (Ulrich 1995). Here, modules here, no matter how diverse they are, are product-dependent. That is, the product is design first and then modules are designed (Yoo et al. 2010, Gulati et al. 2012). In a generative digital ecosystem, where modules are designed first without considering, modules are product agnostic. Here, there is no architecture that provides a boundary of a system. In a way, innovations in digital ecosystems are unbounded (Yoo et al. 2010). In this paper, we offer a new theoretical perspective on innovations in digital ecosystem that do not require a priori architectural knowledge that provides the boundary and interfaces of the system.

Generativity of a Digital Ecosystem

The idea of generativity is a useful analytical concept to understand the dynamic and unbounded aspect of digital innovations (Yoo et al 2010). Zittrain(2006) focus on particular features of digital technology and the behavioral traits of actors in order to explain the generativity of digital artifacts. In this paper, we adopt the (re)combinatorial view of distributed existing software code modules (namely plug-ins) as a way of understanding digital innovations (Chesbrough et al. 2006, Tiwana et al. 2010, Yoo et al. 2010). Specifically, we define the generativity of a digital ecosystem as its reproductive capacity to produce unprompted and uncoordinated changes in its structure and behavior by utilizing existing software modules that can be recombined in novel ways.

At the core of our conceptualization of generativity of a digital ecosystem, there is a rich set of heterogeneous modules that produces an ever-changing landscape of the ecosystem. Many successful digital ecosystems, such as Facebook, Twitter or WordPress, critically depend on the innovations by uncoordinated, heterogeneous third-party developers who create new digital artifacts (such as software plug-in program or Apps) in pursuit of their self-interests and ideas. However, these generative actions by third-party developers are not entirely random as they often rely on existing software modules in building their own digital artifacts. Some of these existing modules that are used in creating new digital artifacts are developed by the original platform owners, but many others are designed by other firms. Therefore, the generativity of a digital ecosystem is an outcome of purposive redefinition

and re-combination of existing heterogeneous modules (Chesbrough et al. 2006, Yoo et al. 2010). Incredibly rich sets of digital artifacts are derived from the simultaneous interactions among pre-existing modules. Therefore, a new breed of digital artifacts in a digital ecosystem can be created through creative recombination of existing modules.

To explain the generativity of a digital ecosystem, we adopt a systems biology perspective that describes how the same sequence of DNA can produce many different cells that lead to diverse functions (Kauffman, 1993).

Theory

Systems Biology Approach

One of the principle goals of systems biology is to understand the generativity of genes in a cell that can produce different functional types of cells. The dynamic gene interactions are derived from the large number of functionally diverse, and frequently multifunctional, sets of elements that interact selectively and nonlinearly to produce coherent functions of a cell (Kitano 2002b). A gene network determines how different genes are expressed in phenotypes through developmental process such as cell differentiation and morphogenesis. A gene is a basic functional unit of the genome of which information stored in DNA is transcribed into mRNA that eventually translates into a protein. A protein is a basic functional and structural unit of a cell. Proteins exhibits coherent expression patterns from the network of diverse genes that they derive from as they interact together to express a specific function. Thus, a gene network plays an essential role in the process by which genotypes give rise to phenotype, determining how different cells perform different functions within an organism although they all share the same genes (Stroagatz 2001, Stuart et al. 2003, Barabasi and Oltvai 2004, Wagner et al. 2007, Ruan et al. 2010).

In particular, a gene co-expression network describes the coherent expression pattern of genes in a cell. A gene co-expression network aims at identifying how subsets of genes are connected among one another. We can think of measuring network connectivity to understand the co-expression patterns. Pearson correlation is widely used in gene expression network by measuring the strength of linear dependence between pair of genes in each cell (Stuart et al. 2003). From functional correlation, a gene co-expression network attempts to represent neighboring relationships (Ruan et al.2010). The neighboring relationship is to include the information about interactions of indirectly correlated genes in a cell (Zhang and Horvath 2005). Based on the neighboring relationship from the correlation value, we can operationalize the topological overlap matrix to find the structure of co-expressed genes (Ravasz et al. 2002, Zhang and Horvath 2005, Horvath 2011).

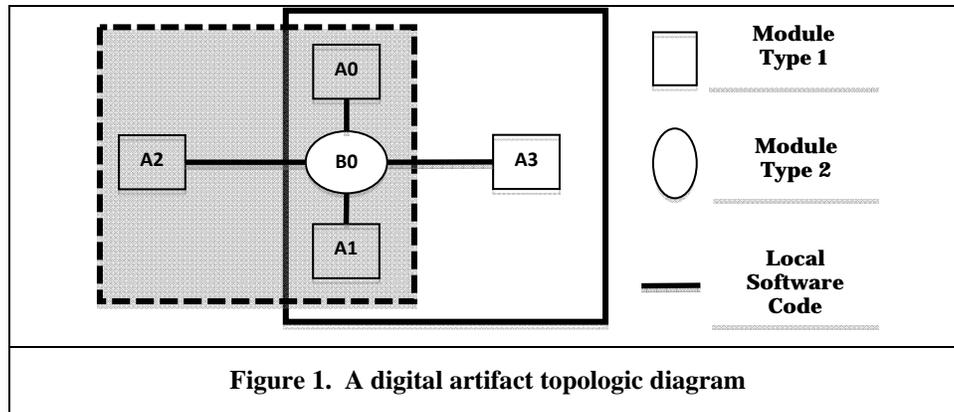
A topology is a structural part of a network. A topological model is a bottom-up approach to find genes that form certain structures (i.e., the connections among genes) (Albert and Othmer 2003). A topology is useful in understanding the large structure of network that is characterized depending on if the network shows random or not (Jeong et al. 2000). Therefore, a topological model can indicate the presence of recurring patterns of gene interactions in a network (Zao et al. 2010). However, a function of a cell is not expressed from a single gene but several genes at the same time, as genes in a cell are expressed together for a single function. To understand the coexpressed network, topological overlap of networks is used (Stuart et al. 2003). Ravasz et al. (2002) explains the meaning of the topological overlap in the broad sense of network approach as the identical part of the large scale organization interconnected one another. In such a way, we can identify the structure of a gene co-expression network from the level of concordance among interactions of genes.

A Network Model of Generativity in a Digital Ecosystems

A digital artifact, such as plug-ins or mobile app, is consisted of existing software modules and software codes that glue these modules together and add additional functionalities such as user interface. Drawing on a systems biology perspective, interactions among modules can be conceptualized as a co-expression network, as various modules used on a digital artifact are coexpressed with one another to have coherent function of a digital artifact. In this sense, the coexpression network gives insight how modules used on digital artifacts are connected with one another. The connectivity can be simplified with the understanding of topological overlap of modules.

Based on the understanding of topological overlap, we can construct a co-expression network model of digital artifacts as shown in Figure 1. In the figure, we have three digital artifacts that are created through combinations of 5 existing modules. Three digital artifacts are: one with a solid box (consisting of modules A0, A1, A3, and B), another one with a dotted box (consisting of modules A0, A1, A2, and B), and the last one in the topologically

overlapped box (consisting of modules A0, A1, and B). The modules are from two different topologies (A's and B). These modules are glued together by local software codes (straight lines connecting modules). The network of A0, B0 and A1 form the basis that two other functionally differentiated two additional digital artifacts through (re)combinations of additional modules, A2 and A3. Each combination of the networks of the modules co-expressed together for the different functions of the digital artifacts.



The main goal of coexpression network model of digital artifact is to explore if there is an underlying design principle among uncoordinated developers in a digital ecosystem by seeking the connectivity of digital artifacts. The underlying structure can be understood once the degree of distribution is specified (Newman 2001). The coexpression network model gives us the possibility of two distinct structural differences between randomness and order, which are specified based on the degree of connectivity of skewedness (Jeong et al. 2000, Watts 2003). Random network model assumes that most nodes have similar number of connectivity so that the distribution is based on Poisson distribution, $P(k) \approx e^{-k}$, where k indicates interaction of modules in this case. In this sense, highly connected modules will show exponential shape. On the other hand, ordered network model assumes that a few nodes are highly connected so that the distribution follows Power law, $P(k) \approx k^{-\gamma}$, where γ follows a particular constant value such as the range from 2.2 to 2.6 in scale free network.

The possibility of two different network structures comes from the uncoordinated developers who can show the different degree of distributed pattern in selecting modules in a digital ecosystem. The module interactions in a digital ecosystem make hard to understand an underlying pattern of generativity. To simplify the interactions, we use topological model drawn on the Figure 1. Using the model, we can explore how different digital artifacts are composed of and share certain modules with another digital artifact (Strogatz 2001, Ravasz et al. 2002). Modules commonly used on each topology can be understood as the basic generative momentum for generativity. Therefore, we expect that the network structure of the topological overlap affect generativity. If there is patterns of generativity in a digital ecosystem, then we can predict how generativity emerges in the next step. Thus, we hypothesize:

H1: There is an underlying structure of generativity in a digital ecosystem

An Empirical Study

Data

We test our hypothesis using the data we collected from WordPress. WordPress, the leading blog engine, is an excellent example of how generativity occurs through the participation of third party developers in an open digital ecosystem. WordPress has a digital ecosystem with over 20,000 plug-ins that enhance the functionality of the original WordPress platform. Most of these plug-ins were developed by third-party developers by connecting external web services such as Google, Facebook and Twitter to a limited set of internal functions that WordPress platform provides. The interactions between the original WordPress functions and external Web services form the “genetic basis” of an incredibly rich array of plug-ins. We attempt to understand the interaction between WordPress functions and external web services through a network biology perspective.

We collected latest version of available 13,491 plug-ins' PHP code data from Wordpress.org. We used a text-mining program using Java script to identify the APIs that are used for each plug-in. We identified APIs in two different ways. For the internal APIs provided by WordPress, we could extract 93 internal APIs provided by the WordPress. For the external APIs, we used Programmable Web to identify 915 web services API from various firms such as Google, Facebook, and Twitter. In such a way, we could construct the data about the network of APIs on WordPress plug-in ecosystem. We constructed 13,491 plug-ins \times 1008 APIs (93 internal APIs + 915 external APIs) matrix to know the network of each API to capture the interaction of APIs among plug-ins in the digital ecosystem.

Empirical Result

To understand the topological pattern based on the scale free topology of 13,491 plug-ins, we first plot a dendrogram, a phylogenetic tool used in biology to help describe the arrangement of topologies. We can generate dendrograms based on the topology of APIs. In figure 2, a dendrogram shows how 1008 APIs that are used to make 13,491 plug-ins are categorized based on the topology. The height in y-axis indicates the combination similarity based on APIs usage that ranges from 0 to 1. The lower value indicates the low similarity. APIs with high interactions are located at the tip of the branches of the dendrogram that indicates the highest interactions comparing to the other APIs in the topology. There are seven largely categorized topologies colored by blue, brown, green, grey, red, turquoise, and yellow. Consecutively, each color is related to 74, 68, 45, 2, 36, 202, and 55 of different topological branches.

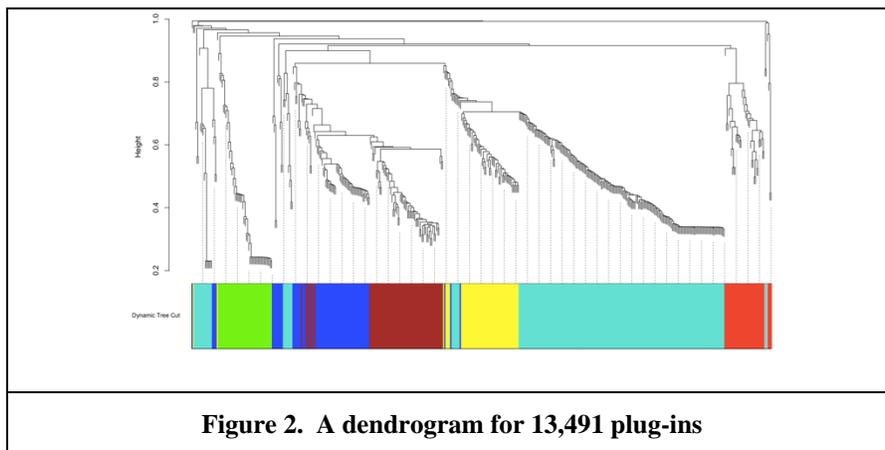


Figure 2. A dendrogram for 13,491 plug-ins

Ravasz et al.(2002) indicates the underlying structure of a complex network by using the topological overlap corresponding to similarity measure of each network group. We construct an $N \times N$ matrix of plugins from plug-ins \times APIs matrix where the different cells are color-coded to represent the extent of overlap of those APIs in the plugin system. The different colors indicate the degree of topological overlap among topologies. In the topological overlap plot, red represents maximum overlap, i.e., if those APIs are shared by many different plugins, and yellow represents if those APIs are not widely shared in any plugin. We can also claim that red is highly overlapped in the network of which the API shows relatively high network centrality to the other.

For the analysis, we adopted the weighted gene co-expression network analysis, which is to analyze DNA microarray data such as brain cancer and mouse genetics in the systems biology Horvath (2011). The analysis allows us to establish the topologies of correlated APIs distributed over a digital (Ravasz et al. 2002). However, because of computational limits to analyzing the entire 13,491 plug-in data set, we randomly sampled 15% of the plug-ins. This subset was the maximum number of plug-ins possible to draw on the topological overlap plot.

In the topological overlap plot, each row and column corresponds to APIs. Higher topological overlap is indicated by darker red, while low topological overlap is shown as light yellow color. Squares correspond to interactions of each APIs on each row and column. Clearly, there are yellow and red colored squares that mean topologically overlapped parts. Different colors within a topology indicate significantly different density among API interactions. The red color indicates the dense part within a topology, while the yellow color indicates the sparse part. Therefore, the different density enables us to identify the relationship among APIs. If the distribution of APIs used on each plug-in is random, there will be no particular red colored topological overlapped building blocks on the figure due to

similar number of connectivity on each module. Thus, the result indicates that there is structural non-random pattern of generativity. So we can understand that there is an underlying structure of generativity in a digital ecosystem by accepting Hypothesis 1.

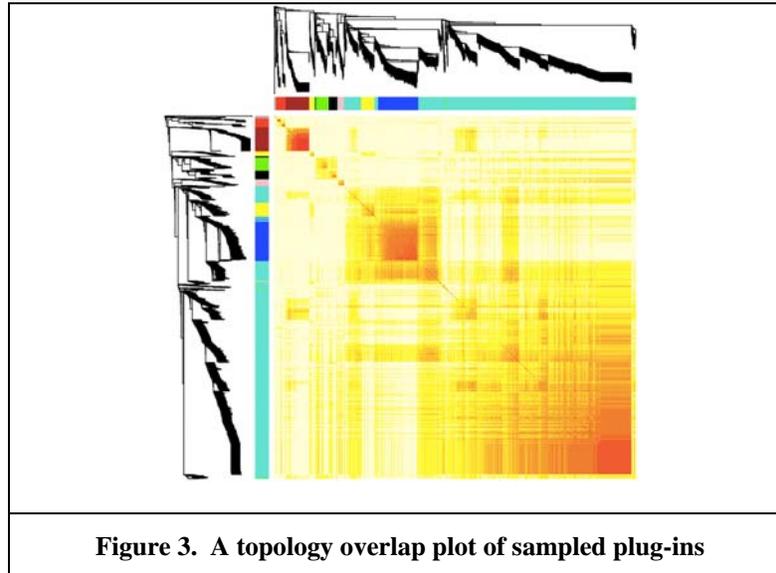


Figure 3. A topology overlap plot of sampled plug-ins

Conclusions

This paper seeks to enhance the understanding of generativity and innovation in relationship with each other by focusing on the structure of APIs. An underlying mechanism of innovation can be understood in two different ways. The analysis of topological overlap provides us information about the interactions among the internal and external APIs by showing which structure is more generative than others in leading to different types of offspring. Thus, the understanding of topological overlap provides a new perspective on innovation that expands our understanding of it.

There are a number of limitations in this study. Though each plug-in API shows heterogenous traits, we cannot normalize them in a systematic way. This limitation also leads to the fact that finding more dynamic innovation patterns from plug-in APIs to understand the interactions of platform owners and third party developers may not be possible.

This research, however, provides a useful guideline to those seeking to understand innovation in a digital ecosystem. By providing the pattern of innovation in both a micro and macro perspective, this study provides a guideline for understanding how continuous innovation can be pursued in an organization under a dynamic environment. This insight indicates how individuals in organizations may have to do away with their ordinary routine work for progress and how managers may need to arrange individual changes by proposing new directions of innovation.

References

- Albert, R. and H. G. Othmer (2003). "The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*." *Journal of theoretical biology* 223(1): 1-18.
- Andrade, R. F. S., J. G. V. Miranda, et al. (2008). "Measuring distances between complex networks." *Physics Letters A* 372(32): 5265-5269.
- Barabási, A. L. and R. Albert (1999). "Emergence of scaling in random networks." *Science* 286(5439): 509-512.
- Barabási, A. L. and Z. N. Oltvai (2004). "Network biology: understanding the cell's functional organization." *Nature Reviews Genetics* 5(2): 101-113.

- Barabási, A. L. (2011). "The network takeover." *Nature Physics* 8(1): 14-16.
- Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*, Yale Univ Pr.
- Boccaletti, S., V. Latora, et al. (2006). "Complex networks: Structure and dynamics." *Physics reports* 424(4): 175-308.
- Boland, R. J., K. Lyytinen, et al. (2007). "Wakes of innovation in project networks: The case of digital 3-D representations in architecture, engineering, and construction." *Organization Science* 18(4): 631-647.
- Baldwin, C. Y. (2008). "Where do transactions come from? Modularity, transactions, and the boundaries of firms." *Industrial and Corporate Change* 17(1): 155-195.
- Baldwin, C. Y. and K. B. Clark (2000). "Design rules, Vol. 1: The power of modularity."
- Bornholdt, S. (2005). "Systems biology: less is more in modeling large genetic networks." *Science's STKE* 310(5747): 449.
- Cabigiosu, A. and A. Camuffo (2011). "Beyond the "Mirroring" Hypothesis: Product Modularity and Interorganizational Relations in the Air Conditioning Industry." *Organization Science*.
- Chesbrough, H. W., W. Vanhaverbeke, et al. (2006). *Open innovation: Researching a new paradigm*, Oxford University Press, USA.
- Clark, K. B. (1985). "The interaction of design hierarchies and market concepts in technological evolution* 1." *Research Policy* 14(5): 235-251.
- Ceci, F. and A. Masini (2011). "Balancing specialized and generic capabilities in the provision of integrated solutions." *Industrial and Corporate Change* 20(1): 91-131.
- Colfer, L. and C. Baldwin (2010). "The mirroring hypothesis: Theory, evidence and exceptions." *Working paper 10-058, Harvard Business School*.
- Erdős, P. and A. Rényi (1960). *On the evolution of random graphs*, Citeseer.
- Eisenmann, T., G. Parker, et al. (2011). "Platform envelopment." *Strategic Management Journal*.
- Eppinger, S. D., D. E. Whitney, et al. (1994). "A model-based method for organizing tasks in product development." *Research in Engineering Design* 6(1): 1-13.
- Erikson, E. H. (1993). *Childhood and society*, WW Norton & Company.
- Fuller, T. F., A. Ghazalpour, et al. (2007). "Weighted gene coexpression network analysis strategies applied to mouse weight." *Mammalian Genome* 18(6): 463-472.
- Gao, L. S. and B. Iyer (2006). "Analyzing complementarities using software stacks for software industry acquisitions." *Journal of Management Information Systems* 23(2): 119-147.
- Garud, R. and A. Kumaraswamy (1993). "Changing competitive dynamics in network industries: An exploration of Sun Microsystems' open systems strategy." *Strategic Management Journal* 14(5): 351-369.
- Geard, N. (2004). "Modelling gene regulatory networks: systems biology to complex systems."
- Geard, N. and K. Willadsen (2009). "Dynamical approaches to modeling developmental gene regulatory networks." *Birth Defects Research Part C: Embryo Today: Reviews* 87(2): 131-142.
- Ghazalpour, A., S. Doss, et al. (2006). "Integrating genetic and network analysis to characterize genes related to mouse weight." *PLoS Genetics* 2(8): e130.
- Gulati, R., P. Puranam, et al. (2012). "Meta-organization design: Rethinking design in interorganizational and community contexts." *Strategic Management Journal* 33(6): 571-586.
- Henderson, R. M. and K. B. Clark (1990). "Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms." *Administrative Science Quarterly*: 9-30.
- Horvath, S. and J. Dong (2008). "Geometric interpretation of gene coexpression network analysis." *PLoS computational biology* 4(8): e1000117.
- Horvath, S. (2011). *Weighted Network Analysis: Applications in Genomics and Systems Biology*, Springer Verlag.
- Jarvenpaa, S. L. and A. Majchrzak (2008). "Knowledge collaboration among professionals protecting national security: Role of transactive memories in ego-centered knowledge networks." *Organization Science* 19(2): 260-276.
- Jeong, H., B. Tombor, et al. (2000). "The large-scale organization of metabolic networks." *Nature* 407(6804): 651-654.
- Kallinikos, J. and J. C. Mariátegui (2011). "Video as digital object: production and distribution of video content in the internet media ecosystem." *The information society* 27(5): 281-294.
- Kaufman, L. and P. J. Rousseeuw (1990). *Finding groups in data: an introduction to cluster analysis*, Wiley Online Library.
- Kauffman, S. A. (1993). *The origins of order*, Oxford University Press New York.
- Kitano, H. (2002a). "Systems biology: a brief overview." *Science* 295(5560): 1662-1664.
- Kitano, H. (2002b). "Computational systems biology." *Nature* 420(6912): 206-210.

- Kulldorff, M. (1997). "A spatial scan statistic." *Communications in statistics-theory and methods* 26(6): 1481-1496.
- Langlois, R. N. (2003). "The vanishing hand: the changing dynamics of industrial capitalism." *Industrial and Corporate Change* 12(2): 351-385.
- Lee, M. T. and J. Kim (2008). "Self containment, a property of modular RNA structures, distinguishes microRNAs." *PLoS computational biology* 4(8): e1000150.
- MacCormack, A., C. Baldwin, et al. (2010). "The Architecture of Complex Systems: Do Core-periphery Structures Dominate?" *MIT Slone Research Paper no. 4770-10*.
- Mayo, O. and R. Buerger (1997). "The evolution of dominance: a theory whose time has passed?" *Biological Reviews* 72(1): 97-110.
- McAdams, D. P. and E. de St Aubin (1992). "A theory of generativity and its assessment through self-report, behavioral acts, and narrative themes in autobiography." *Journal of Personality and Social Psychology* 62(6): 1003.
- MacCormack, A. D., J. Rusnak, et al. (2008). Exploring the duality between product and organizational architectures: A test of the mirroring hypothesis, Harvard Business School.
- Newman, M. E. J., S. H. Strogatz, et al. (2001). "Random graphs with arbitrary degree distributions and their applications." *Physical Review E* 64(2): 026118.
- Newman, M. E. J., D. J. Watts, et al. (2002). "Random graph models of social networks." *Proceedings of the National Academy of Sciences of the United States of America* 99(Suppl 1): 2566-2572.
- Provan, K. G., A. Fish, et al. (2007). "Interorganizational networks at the network level: A review of the empirical literature on whole networks." *Journal of management* 33(3): 479-516.
- Ravasz, E., A. L. Somera, et al. (2002). "Hierarchical organization of modularity in metabolic networks." *Science* 297(5586): 1551-1555.
- Ravasz, E. and A. L. Barabási (2003). "Hierarchical organization in complex networks." *Physical Review E* 67(2): 026112.
- Ruan, J., A. K. Dean, et al. (2010). "A general co-expression network-based approach to gene expression analysis: comparison and applications." *BMC systems biology* 4(1): 8.
- Salton, G., A. Wong, et al. (1975). "A vector space model for automatic indexing." *Communications of the ACM* 18(11): 613-620.
- Sambamurthy, V., A. Bharadwaj, et al. (2003). "Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms." *Mis Quarterly*: 237-263.
- Sanchez, R. and J. T. Mahoney (1996). "Modularity, Flexibility, and Knowledge Management in Product and Organization Design." *Strategic Management Journal*: 63-76.
- Sanchez, R. (2004a). "Understanding competence-based management: Identifying and managing five modes of competence." *Journal of Business research* 57(5): 518-532.
- Sanchez, R. (2004b). "Creating modular platforms for strategic flexibility." *Design Management Review* 15(1): 58-67.
- Sanchez, R. (2004c). "Understanding competence-based management: Identifying and managing five modes of competence." *Journal of Business research* 57(5): 518-532.
- Schilling, M. A. (2002). "Modularity in multiple disciplines." *Managing in the Modular Age: Architectures, Networks and Organizations*: 203-214.
- Schumpeter, J. A. (1943). "Capitalism, socialism and democracy."
- Simon, H. A. (2002). "Near decomposability and the speed of evolution." *Industrial and Corporate Change* 11(3): 587-599.
- Spender, J. C. and P. H. Grinyer (1995). "Organizational renewal: Top management's role in a loosely coupled system." *Human Relations* 48(8): 909-926.
- Sosa, M. E., S. D. Eppinger, et al. (2003). "Identifying modular and integrative systems and their impact on design team interactions." *Journal of Mechanical Design* 125: 240.
- Sosa, M. E., S. D. Eppinger, et al. (2004). "The misalignment of product architecture and organizational structure in complex product development." *Management science*: 1674-1689.
- Sosa, M. E., S. D. Eppinger, et al. (2007). "A network approach to define modularity of components in complex products." *Journal of Mechanical Design* 129: 1118.
- Staudenmayer, N., M. Tripsas, et al. (2005). "Interfirm Modularity and Its Implications for Product Development*." *Journal of Product Innovation Management* 22(4): 303-321.
- Stone, R. B., K. L. Wood, et al. (2000). "A heuristic method for identifying modules for product architectures." *Design studies* 21(1): 5-31.
- Strogatz, S. H. (2001). "Exploring complex networks." *Nature* 410(6825): 268-276.

- Stuart, J. M., E. Segal, et al. (2003). "A gene-coexpression network for global discovery of conserved genetic modules." Science 302(5643): 249-255.
- Swanson, E. B. (1994). "Information systems innovation among organizations." Management science: 1069-1092.
- Tegner, J., M. Yeung, et al. (2003). "Reverse engineering gene networks: integrating genetic perturbations with dynamical modeling." Proceedings of the National Academy of Sciences 100(10): 5944.
- Tiwana, A. (2008). "Does technological modularity substitute for control? A study of alliance performance in software outsourcing." Strategic Management Journal 29(7): 769-780.
- Tiwana, A., B. Konsynski, et al. (2010). "Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics." Information Systems Research 21(4): 675-687.
- Tushman, M. L. and L. Rosenkopf (1992). "Organizational determinants of technological change: toward a sociology of technological evolution." Research in organizational behavior 14: 311-347.
- Ulrich, K. (1995). "The role of product architecture in the manufacturing firm." Research Policy 24(3): 419-440.
- Ulrich, K. T. and S. D. Eppinger (2011). Product design and development, McGraw-Hill.
- Unger, D. and S. Eppinger (2011). "Improving product development process design: a method for managing information flows, risks, and iterations." Journal of Engineering Design 22(10): 689-699.
- Wagner, A. and D. A. Fell (2001). "The small world inside large metabolic networks." Proceedings of the Royal Society of London. Series B: Biological Sciences 268(1478): 1803-1810.
- Wagner, G. P., M. Pavlicev, et al. (2007). "The road to modularity." Nature Reviews Genetics 8(12): 921-931.
- Wang, B., J. M. Phillips, et al. (2008). Spatial scan statistics for graph clustering.
- Watts, D. J. (2003). Small worlds: the dynamics of networks between order and randomness, Princeton university press.
- Whitney, D. E. (2004). Mechanical Assemblies-Their Design, Manufacture, and Role in Product Development, New York: Oxford Series on Advanced Manufacturing.
- Yoo, Y., O. Henfridsson, et al. (2010). "Research Commentary---The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research." Information Systems Research 21(4): 724-735.
- Zhao, W., P. Langfelder, et al. (2010). "Weighted gene coexpression network analysis: state of the art." Journal of Biopharmaceutical Statistics 20(2): 281-300.
- Zhang, B. and S. Horvath (2005). "A general framework for weighted gene co-expression network analysis." Statistical applications in genetics and molecular biology 4(1): 1128.
- Zittrain, J. L. (2006). "The generative internet." Harvard Law Review: 1974-2040.