6

Understanding Business Ecosystem Dynamics: A Data-Driven Approach

RAHUL C. BASOLE, Georgia Institute of Technology
MARTHA G. RUSSELL, Stanford University
JUKKA HUHTAMÄKI, Tampere University of Technology
NEIL RUBENS, University of Electro-Communications
KAISA STILL, VTT Technical Centre of Finland
HYUNWOO PARK, Georgia Institute of Technology

Business ecosystems consist of a heterogeneous and continuously evolving set of entities that are interconnected through a complex, global network of relationships. However, there is no well-established methodology to study the dynamics of this network. Traditional approaches have primarily utilized a single source of data of relatively established firms; however, these approaches ignore the vast number of relevant activities that often occur at the individual and entrepreneurial levels. We argue that a data-driven visualization approach, using both institutionally and socially curated datasets, can provide important complementary, triangulated explanatory insights into the dynamics of interorganizational networks in general and business ecosystems in particular. We develop novel visualization layouts to help decision makers systemically identify and compare ecosystems. Using traditionally disconnected data sources on deals and alliance relationships (DARs), executive and funding relationships (EFRs), and public opinion and discourse (POD), we empirically illustrate our data-driven method of data triangulation and visualization techniques through three cases in the mobile industry Google's acquisition of Motorola Mobility, the coopetitive relation between Apple and Samsung, and the strategic partnership between Nokia and Microsoft. The article concludes with implications and future research opportunities.

Categories and Subject Descriptors: H.2.8 [**Database Management**]: Database Applications—*Data mining*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Graphical user interfaces*; I.2.7 [**Artificial Intelligence**]: Natural Language Processing—*Text analysis*

General Terms: Design, Measurement, Performance

Additional Key Words and Phrases: Data triangulation, information visualization, interorganizational networks, business ecosystem

ACM Reference Format:
Rahul C. Basole, Martha G. Russell, Jukka Huhtamäki, Neil Rubens, Kaisa Still, and Hyunwoo Park. 2015. Understanding business ecosystem dynamics: A data-driven approach. ACM Trans. Manag. Inform. Syst. 6, 2, Article 6 (May 2015), 32 pages.
DOI: http://dx.doi.org/10.1145/2724730

Authors' addresses: R. C. Basole, Georgia Institute of Technology, School of Interacting Computing & Tennenbaum Institute, Atlanta, Georgia 30332 USA; email: basole@gatech.edu; M. G. Russell, mediaX at Stanford University, Stanford, California 94305-4115, USA; email: martha.russell@stanford.edu; J. Huhtamäki, Tampere University of Technology, Department of Mathematics, Korkeakoulunkatu 3, FI-33720 Tampere, Finland; email: jukka.huhtamaki@tut.fi; N. Rubens, Graduate School of Information Systems, University of Electro-Communications, Tokyo 182-8585, Japan; email: rubens@ai.is.uec.ac.jp; K. Still, VTT Technical Research Centre of Finland Ltd, Kaitoväylä 1, FI-90571 Oulu, Finland; email: kaisa.still@vtt.fi; H. Park, Georgia Institute of Technology, School of Industrial and Systems Engineering & Tennenbaum Institute, 755 Ferst Drive, Atlanta, Georgia 30332; email: hwpark@gatech.edu.

DOI: http://dx.doi.org/10.1145/2724730

1. INTRODUCTION

Business ecosystems consist of a heterogeneous and continuously evolving set of individuals and firms that are interconnected through a complex, global network of relationships. These firms come from a variety of market segments, each providing unique value propositions [Basole 2009]. It is quite unlikely for a single market segment to deliver all products or services to end-consumers. In fact, value creation and delivery require a careful orchestration between firms across segments [Dhanaraj and Parkhe 2006]. In the mobile ecosystem, for instance, the massive rollouts and upgrades of cellular networks by mobile network operators would be useless without devices that can fully leverage them. Similarly, smartphones would just be boxes with little or no value without a platform and platform-enabled applications. App stores provide third-party developers ways to offer content and reach consumers. Cocreation is hence an essential ecosystem characteristic, because a continual realignment of synergistic relationships of knowledge, resources, and talent is required for growth of the system and responsiveness to changing internal and external forces [Rubens et al. 2011].

However, methodological approaches to comprehensively study the dynamics of business ecosystems are limited [Ahuja et al. 2011]. Existing approaches have usually focused on events and activities at singular levels, with prominent examples in the biotechnology [Owen-Smith and Powell 2004], software [Iyer et al. 2006], and information technology industries [Iansiti and Richards 2006]. A methodology that would more comprehensively describe how an ecosystem emerges and evolves depends on data from multiple sources. Several potentially complementary data sources containing relevant stakeholder activity information exist some proprietary, others publicly available and still emerging (such as social media). Historically, data acquisition for business ecosystem research was a resource-intensive step. Open access to online data has made a wealth of data widely available and permits researchers to leverage it for insights about the emergence and evolution of ecosystems. A key challenge has now become the choice and qualification of data for analysis. The different data sources are often disconnected and reflect different units and periodicity; they are rarely interoperable. In some instances there is overlap, in others they are complementary, and in others they provide different and even conflicting insights.

Our article contributes in several ways. Methodologically, our article introduces the use of different but complementary datasets, contributing to the understanding of how large, disconnected, potentially complementary structured and unstructured datasets can be used concurrently in an attempt to provide multilevel base-specific details at the same time as revealing the big-picture overview perspectives. It also presents the use of custom network visualizations and small multiples to reveal concurrent patterns of different ecosystem relationships. Our objective in implementing this methodology is to activate exploration and discovery and to generate insights on how business ecosystem evolution can be visually represented and understood.

Managerially, our article provides competitive intelligence and insights into the systemic behavior and outcomes of firms, thus enabling the “wide lens” [Adner 2012; Basole 2014]. Small companies want to identify opportunities. Large companies want to know what’s around the corner regarding technology and innovation, about which competitors they should worry, and with which collaborators they should partner. Large or small, they want to learn who has succeeded, why, and how long it took. The method and tools introduced in this article provide such capabilities.

The remainder of the article is structured as follows. Section 2 reviews the related work. Section 3 describes our methodology, data sources, and analysis. In Section 4, we present three illustrative examples. Implications are discussed in Section 5. Section 6 concludes the article and presents opportunities for future research.

2. RELATED WORK

Our article draws on four distinct but interrelated literature streams: interfirm networks and ecosystems, socially constructed and curated data, data triangulation, and visualization and visual analytics.

2.1. Interfirm Networks and Business Ecosystems

A business ecosystem consists of interdependent (1) firms that form symbiotic relationships to create and deliver products and services [Basole and Rouse 2008; Dougherty and Dunne 2011] integrated with (2) human actors that form human networks, such as linkages between executives [Hwang and Horowitt 2012; Ibarra and Hansen 2011; Vargo and Lusch 2004]. The conceptualization of markets as ecosystems is a result of theoretical extensions of research on inventor networks [Powell and Giannella 2009] and of interfirm networks, alliances, and innovation [Gulati 1998; Moore 1993; Oliver 1990]. As product and service development have become increasingly disintegrated vertically and horizontally, the imperative need for creating interfirm relations has been noted [Iansiti and Levien 2004]. The formation of networks and alliances has been particularly beneficial in technology industries as it has allowed firms to share risks in development, to have access to synergistic knowledge [Eisenhardt and Schoonhoven 1996], and to be responsive to change in the external environment [Russell et al. 2011]. Studies have shown that interfirm networks are an effective organizational form to improve firm performance, speed of innovation, and organizational learning [Ahuja 2000; Gulati et al. 2000].

The business ecosystem structure reflects the interactions formed by complex networks [Green and Sadedin 2005]. Studies have adopted a complex networked systems perspective to examine why, when, and how interfirm networks and alliances form and change [Ahuja 2000; Gulati et al. 2000]. This view combines both the resource dependency and embeddedness perspectives and suggests that interfirm networks are complex systems characterized by coevolving actors engaged in collaboration, coopetition [Iansiti and Levien 2004], and collective invention [Powell and Giannella 2009].

2.2. Provenance in Socially Constructed and Curated Data

In the current tsunami of digital business data, the provenance of this data is of critical importance. The process of data curation seeks to ensure that data is continuously available and fit for discovery, analysis, and reuse [Lord and Macdonald 2003]. Curated data is perceived to have the advantages of consistent ontologies, predictable data-gathering methods, and consistently applied data-cleaning rules. With the standardized data practices and policies of curated data, analytical methods can become standardized, and interpretation of analytical results benefits from consistent comparisons and a shared understanding of metrics. These very advantages, however, bring with them some disadvantages. Bias becomes baked into data policies, and only parts of data get curated. Categories and classifications sometimes persist in data practice long after real-world semantics have shifted to new classifications or reformulated categories. The time required for the curation processes may introduce significant delays into the timeliness of even the most recently available curated data. Additionally, many curated databases have limited availability and access may be exclusive and/or very expensive.

While some have argued that data in and of itself has little meaning and that the knowledge [Borgman 2007] and meaning of data [Smagorinsky 1995] are inherently socially constructed, the social nature of the Internet has added a new data frontier in socially constructed data. Extensive data about businesses is now openly available through company websites, published announcements and filings, blogposts and microblogging, and community-built information resources. These sources provide unprecedented access to data, updated in real time, and also include individuals' activities and interactions. One of the first of its kind, Wikipedia established itself as the most reliable source of accurate information [Giles 2005] because it invited additions and tracked the provenance of changes; a data source that is socially constructed has observable patterns of governance [Leskovec et al. 2010].

2.3. Data Triangulation

While data curation improves the quality and accessibility of data, a significant risk for acceptance of data-driven ecosystem research lies with defending the validity of data and its limitations. If the data is incorrect or is not credible, the insights gained and conclusions drawn can be misleading. Indeed, validity is a major threat for data-driven research [Barnes and Vidgen 2006]. It has been argued that data triangulation not only provides a much richer, complete picture of the phenomena under investigation but also validates, cross-checks, and strengthens findings, in particular when data from different sources point to congruent insights [Kaplan and Duchon 1988].

The concept of triangulation has its origin in navigation, military strategy, and surveying, where it was used to determine positions of items and landmarks by using two or more locations of other items [Hammersley and Atkinson 1983; Jick 1979; Smith 1975]. Denzin [2009] defines data triangulation as the process of using a variety of data sources in the same study. While data triangulation is a resource-intensive task, the benefits are multifold. It permits researchers to be more certain of their findings, can unravel contradictions, and can even lead to the fusion of theories; it can also serve as a litmus test for competing theories and even help in the understanding of processes over time [Bizzi and Langley 2012; Jick 1979; Onwuegbuzie and Leech 2007].

Triangulation is particularly valuable in complex research contexts such as the study of businesses, markets, and ecosystems, characterized by high data volume, velocity, and variety, where data-driven decisions are deemed particularly valuable [McAfee and Brynjolfsson 2012]. The more data sources are used, the more likely a more adequate representation of the underlying phenomenon can be gained [Newman and Ridenour 1998].

Blaikie [1991] argued that if researchers rely on a single data source, there is the potential danger that undetected error in the data production process may render the analysis incorrect. On the other hand, if multiple diverse data lead to the same conclusion, researchers can be a bit more confident in that conclusion [Hammersley and Atkinson 1983]. Indeed, a key purpose behind using multiple sources is to strengthen validity because multiple sources provide multiple measures of the same phenomenon [Yin 2008] and make possible findings that could not be made using a single data source [Patton 2001]. They also provide multiple perspectives on the same phenomenon.

The use of data triangulation in academic research has thus gained substantial traction over the years [Gallivan 1997; Gibbert et al. 2008; Kaplan and Duchon 1988]. Examples include the use of multiple and different data sources, research methods, theories, and investigators to obtain corroborating evidence [Patton 2001] in contexts such as email use [Markus 1994], e-business [Barnes and Vidgen 2006], software engineering [Bratthall and Jørgensen 2002], business performance [Venkatraman and Ramanujam 1987], and market research [Carson et al. 2001]. In this study, we show that triangulation using multiple data sources is particularly valuable in understanding the structure and dynamics of business ecosystems.

2.4. Visualization and Visual Analytics

Because of the complexity of business ecosystems, the derivation of conceptual insight from data can be particularly challenging [Bizzi and Langley 2012]. While an analytical approach provides valuable insights to the structure and dynamics of ecosystems, important knowledge can also be gained through the visual revelation of patterns in a complex business ecosystem's data [Basole et al. 2013]. Contrary to the perception that visualizations are merely artistic approaches to depicting structure, they can help users to see through the forest of data [Fox and Hendler 2011]. Visualizations have been used to explore, interpret, and communicate data in order to aid humans in overcoming their cognitive limitations, making structure, patterns, relationships, and themes visible and providing a means to efficiently compare multiple representations of data in similar fields such as medicine, dentistry, computer science, and engineering. It has been suggested that visualization approaches can be extremely valuable for understanding and analyzing business issues, including strategy, scenario planning, and problem solving [Tufte 1983]. The science of visual analytics seeks to find ways to support analytical reasoning with interactive visual interfaces [Thomas and Cook 2006]. The use of visualizations is advantageous for establishing a common ground for discussion and for helping individuals see their position in a larger context.

3. METHODOLOGY

3.1. Data

Business ecosystem analysis is enabled by an increasing availability of digital data. Given that our objective is to study the structure and dynamics of the ecosystem, identification of relationship data is key. Our study uses three complementary data sources, namely, deals and alliance relationships (DARs) using SDC Platinum, executive and funding relationships (EFRs) using the IEN Dataset, and public opinion and discourse (POD) using Northern Light.[1] Because the validity of ecosystem analysis insights depends on the nature and quality of the datasets, we first describe the datasets and their complementarity and then explain our proposed methodological approach.

3.1.1. Deals and Alliance Relationships (DARs). To study the publicly disclosed and documented deals and alliance relationships among companies, we use data from SDC Platinum, one of the most prominent, comprehensive, and accurate commercial databases used in the study of global interfirm relationships across multiple sectors [Schilling 2009]. It has been used extensively in strategic management, finance, and the management and organization sciences (e.g., Hsu [2006], Sampson [2004], and Schilling and Phelps [2007]). Strategic alliance relationships are only one aspect of this broad database. It also contains information on joint ventures, R&D agreements, sales and marketing agreements, supply and manufacturing agreements, and licensing and distribution data, curated from SEC filings, trade publications, and wire and news sources. One of the primary advantages of the SDC database is its global and cross-industry scope [Schilling 2009]. It covers at least one alliance of each of the 1,000+ four-digit Standard Industrial Classification (SIC) codes and also includes agreements between universities and government labs. In addition, it provides searchable access to 200+ additional data elements, including names, SIC codes and nationality of participants, and relationship terms and synopses, and is updated on a monthly basis.

The SDC database does have some limitations. One limitation is its primary focus on large, public companies. Many small companies, in particular startups, are not represented in the database. The second limitation of the SDC database is the method of data collection. Firms are not required to report alliances to any governing body (e.g., the SEC). While most firms announce their alliances publicly, reporting can be highly variable. Despite this limitation, however, Schilling [2009] found that across multiple datasets, SDC provided the most complete set of alliance announcements and that coding of interfirm relationships was highly accurate. Lastly, while data exists from 1988 forward, it is somewhat sparse until 1990 [Anand and Khanna 2000]. This

[1] It should be noted that our approach is not limited to just three sources, but rather is scalable to n sources. For instance, we could have also used patent and social media data in our analysis. However, for demonstration purposes, we found adding additional sources to be beyond the scope of this study.

limitation does not apply to our study as our focus is on understanding ecosystem dynamics of more recently emerging industries (i.e., ICT).

3.1.2. Executive and Funding Relationships (EFRs). To study business ecosystems from the perspective of executive and funding relationships, we use the Innovation Ecosystems Network (IEN) Dataset [Rubens et al. 2010]. The IEN Dataset is a quarterly updated collection of socially constructed and curated data about technology-oriented companies in the information communication technology fields and the investors and service companies (legal, accounting, advertising) that support them. It includes data about more than 100,000 companies (including enterprises and startup companies), their executives and board personnel, investment organizations, and venture capital investments. People included in the dataset are key individuals in their respective companies (e.g., founders, executives, lead engineers, etc.), members of boards of advisors, and investors. The dataset further includes data on the background of individuals including the degrees they have received from various educational institutions.

3.1.3. Public Opinion and Discourse (POD). To understand business ecosystems from the standpoint of public opinion and discourse, we use data from Northern Light, a leading provider of strategic research portals, business research content, and search technology to global enterprises. It aggregates and indexes press releases, newsfeeds, and the entire collections of over 130 IT analyst, market research, corporate research, think tank, and technology research firms. Northern Light has developed MI Analyst, a powerful web-accessed software tool that through automatic meaning extraction provides deep insights into a large corpus of these structured and unstructured document repositories [Basole et al. 2013]. Meaning extraction has been shown to dramatically improve and accelerate a searcher's ability to gain insight into a topic and answer specific research questions [Seuss 2009].

The meaning extraction process begins with identification and extraction of meaning-loaded concepts, such as events, conditions, situations, outcomes, actions, and relationships [Seuss 2011]. These concepts are organized into an extensive meaning taxonomy. Patterns of related concepts are interpreted by human experts and organized into scenarios. It has been found, through extensive studies, that concepts within a paragraph of each other (e.g., 40 words) are related [Seuss 2011]. MI Analyst enables users to quickly gain important and summarized insights by identifying concepts and presenting scenarios present in the data [Basole et al. 2013]. In doing so, MI Analyst provides critical knowledge discovery support.

We leverage the strength of the MI Analyst solution in our data-driven analysis of the mobile ecosystem. Specifically, we use the POD data to identify and weigh the presence of key scenarios that occurred before, during, and after a significant change in one or more of the network and node-level properties. In doing so, we are able to present a context and potential evidence to specific events surrounding the change in the network structure, providing us a data-driven way to make sense of our network analysis and allow us to tell a much richer story with data.

3.1.4. The Complementarity of the Datasets. The utilization of the three datasets affords significant complementary value for the analysis of ecosystem dynamics. While the DAR dataset contains institutionally validated alliance information for primarily large, global, and public companies, the EFR dataset contains socially validated information about small, private companies and startups, as well as large enterprises.

Many innovation activities occur in entrepreneurial settings or in the context of executive and stakeholder relationships; the EFR dataset presents a relationship perspective and fills in the blanks between major ecosystem events. In contrast to known-quality DAR data, with time delays and potential editorial bias often extant in traditional data settings, EFR data inherits both the advantages and disadvantages of socially constructed data. Some of the advantages are availability, large coverage, timeliness, and community verification of data quality. Some of the disadvantages are potentially erroneous data and public bias. The POD dataset provides meaning-extracted content, derived from newsfeeds, trade journals, blogs, and analyst reports, which together supports our analysis, explanation, and interpretation and provides context to the relational data generated by the DAR and ERF datasets. A comparative summary of the three datasets is provided in Table I.

3.2. Approach

Given our curated datasets, we propose a four-stage process (see Figure 1) for analyzing the dynamics of business ecosystems, consisting of (1) boundary specification; (2) metrics identification; (3) computation, analysis, and visualization; and (4) sense making and storytelling. Our "human-in-the-loop" approach builds on the well-established information visualization reference model [Card et al. 1999] by carefully balancing data management, visual mappings, computer graphics, and interaction for the purpose of creating a flexible and reusable ecosystem visualization platform.

3.2.1. Step 1: Boundary Specification. Boundary specification involves determining the primitives of the network architecture [Ahuja et al. 2011], including nodes, node types (e.g., firms, people, universities, etc.), and relationship types (e.g., R&D, supply chain, marketing, licensing, investing, etc.) and specifying the desired analysis timeframe (e.g., start/end date). The choice of these parameters is driven by the nature and intent of the problem, the questions being asked, and the costs involved.

The specification of nodes is not a trivial task, as firms continuously enter and leave the ecosystem. If the analytical focus is on the evolution of a particular market segment, one may begin by considering all companies that operate in that market sector and the second-level companies to which the selected first-level companies have documented relationships. This leads to a related decision concerning the number of third, fourth, and subsequent levels of nodes to include in the selected data. Which other companies, individuals, and investors should be included in the dataset: only those directly connected to nodes in the first-level market sector or those connected by k-steps? The upper bound limit of k is defined by the maximum k-steps of the graph. The larger the selected k is, the more entities will be included; this expansion carries risks of diluting the analysis with potentially irrelevant nodes. The smaller the selected k is, the greater the risk of ignoring important companies, people, and investors that may be a few steps removed.

The specification of the appropriate timeframe is an equally challenging task. In an era when data was scarce, researchers often chose the largest timeframe available (e.g., the first activity for any of the companies involved in the alliance). The luxury of big data, however, invites careful consideration of the selection criteria. Data selection criteria can be based on a particularly important or relevant point in time (e.g., announcement, product launch, policy decision), on the timing of events and activities perceived to have led to the alliance, or on other study variables.

3.2.2. Step 2: Metrics Identification. Insight objectives and decision processes guide the selection of metrics for analysis of ecosystem dynamics. There are many social network as well as information- and graph-theoretic metrics that can be useful for understanding the dynamics of an ecosystem. Broadly, these can be categorized at two levels of analysis the network as a whole (ecosystem) and the node level (firm/individual). This differentiation is important because network dynamics at each level, although related, are also distinct [Zaheer et al. 2010]. Based on prior experience, we compute three metrics at the network level (i.e., size, density, and number of components) and

Table I. Comparison of Business Ecosystem Datasets

	Deals and Alliance Relationships (DARs)	Executive & Funding Relationships (EFRs)	Public Opinion & Discourse (POD)
Source	SDC Platinum 4.0	IEN Dataset	Northern Light
Source Details	Proprietary (Thomson Reuters Financial) based on U.S. SEC data	Open source based on socially curated data from news, press releases, and social media	Press releases, newsfeeds, analyst reports, market research, corporate research, reports from think tanks and technology research firms
Ecosystem Entities	Firms	Firms, Investors, Individuals	Firms, Individuals, Technologies
Type of Data	Alliance data (strategic, R&D, marketing, manufacturing, licensing, and supply) and status (active, terminated, pending) of public and private firms (37 SIC Codes, four-digit)	Relationship Data of Public and Private Firms (e.g., SME and startups), Financial Organizations, Educational Institutions, Funding Rounds, Acquisitions, Investments by Individuals and Companies	Full-Text content and indexed entities, including business issues, key technologies, individuals, companies, and venture financed companies
Years Covered	1/1/1990 - 12/31/2012	1/1/1994 - 12/31/2012	1/1/2008-12/31/2012
Update Frequency	Monthly	Quarterly	Hourly

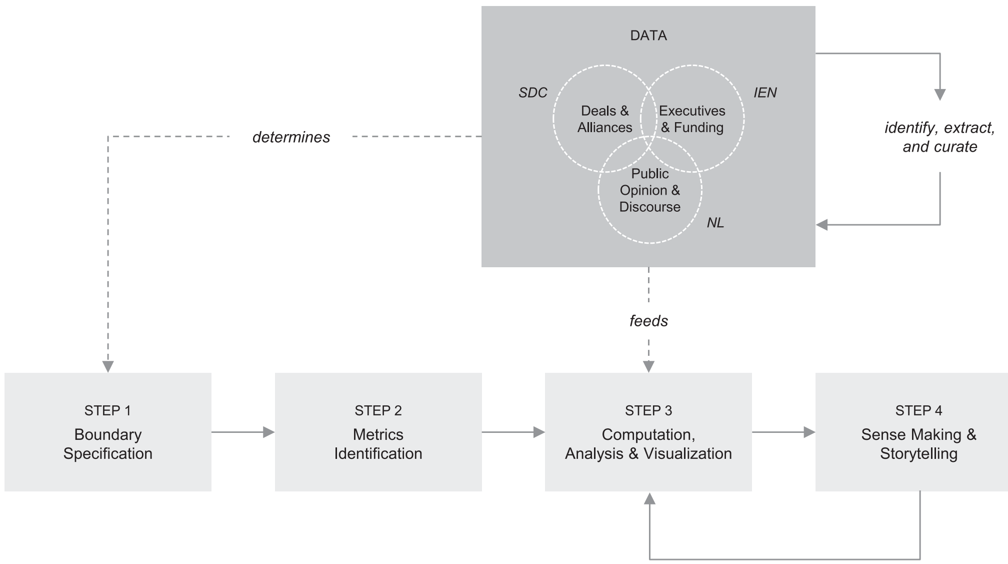

Fig. 1. Four-stage approach to data-driven business ecosystem analysis and visualization.

three metrics at the node level (i.e., degree, betweenness centrality, and clustering coefficient), each at 6-month intervals, to describe the dynamics of business ecosystems [Basole 2009; Basole et al. 2012; Basole 2014]. Size, number of components, and degree reveal insights about the types of links in the ecosystem. Density, betweenness centrality, and clustering coefficient reveal insights about the structure of the ecosystem. The choice of these metrics is driven by the primary motivation to understand firm position, ecosystem cohesion, and clustering. A description of these commonly used node- and network-level metrics [Ahuja et al. 2011] is provided in Table II.

3.2.3. Step 3: Computation, Analysis, and Visualization. There are many ways of analyzing and visualizing temporal, relational ecosystem data [Basole et al. 2012] and many tools that can be used [Huisman and van Duijn 2005]. One approach includes a tabular description of key metrics; another includes a timeline representation of changes in key network metrics. If multiple metrics are to be compared simultaneously and if structural patterns matter more than specific metric levels, sparklines or small multiples are a frequent choice [Tufte 1983]. An interactive, animated visual interface is commonly used in analysis as a means of navigating through alternative views of business ecosystems. An iteration between analysis and visualization sometimes requires a return to data curation and an excursion into sense making to identify a synergistic combination of numerical, visual, and narrative descriptions. Based on prior experiences [Basole et al. 2012, 2013; Basole 2014; Still et al. 2012] and due to page and publication medium constraints, we utilize a combination of tabular representations, cumulative network visualization, and small multiples to depict the dynamics of business ecosystems.

Guided by our objective in understanding the ecosystem dynamics of pairs of focal firms, we developed a novel, custom network visualization layout called the *bicentric ecosystem* layout. This layout uses ideas from set visualization [Alsallakh et al. 2014] and provides an effective representation of two focal firms as well as their shared direct and indirect ecosystem partners. A conceptual representation of this layout is shown in Figure 2. Two firms A and B are placed at a distance d apart; two concentric circles with

Table II. Node and Network-Level Ecosystem Dynamics Metrics

Level	Metric	Description	Formula	Parameter Description
Network	Size	Change in the size of the network is reflective of the overall growth of the relevant ecosystem.		
	Density	Change in density (the proportion of ties that are realized in the network relative to the hypothetical maximum possible) represents how tightly the network is connected.	$DENS = \frac{L}{\frac{N \times (N-1)}{2}}$	where N represents the total number of firms in the ecosystem and L represents the total number of relationships (links) in the ecosystem.
	Number of Components	A component is a subgraph in which any two nodes are connected to each other by paths, and which is connected to no additional nodes in the supergraph.		
Node	Degree	Change in the degree is reflective of the number of new connections a firm has gained or established.	$DEG = n_i$	where n_i represents direct partners of focal firm i
	Betweenness Centrality	Change in betweeness centrality measure is reflective of the positional prominence of a firm (node) in a network.	$BC_i = \sum_{q,i \neq j} p_{q,i,j}$	where $p_{q,i,j}$ represents the proportion of shortest paths between q and j that run through i
	Cluster Coefficient	Change in the cluster coefficient is reflective of the level of connectivity between a firm's directly connected partners.	$CC_i = \frac{n_p}{\frac{n_i \times (n_i - 1)}{2}}$	where t represents the number of existing ties among all n_i direct partners p of focal firm i

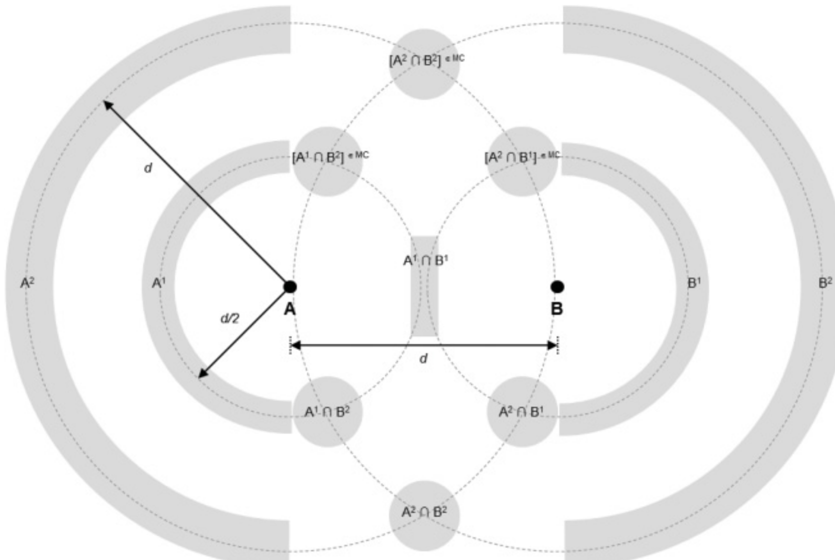

Fig. 2. Conceptual representation of bicentric ecosystem layout.

radii $d/2$ and d, respectively, are drawn around each focal firm. The intersection points of the various circles are areas in which direct (first-tier) and indirect (second-tier) ecosystem partners are placed. Partners shared directly by both focal firms are placed in the middle. Moreover, we differentiate between entities placed at the top and the bottom intersection points. Entities at the top belong to the main component; all others are placed at the bottom. This enables us to further understand the cohesion pattern of subtier levels. To reduce clutter and improve aesthetics, we apply a "no overlap" and node jittering rule. Furthermore, we position nodes within each area by node type for sources with multiple node entities (i.e., firm, people, and investors). This design rule allows us to identify the extent of node type composition. A pseudocode of our bicentric ecosystem layout is provided in Algorithm 1.

Our other objective was to visualize the entire business ecosystem. Our second visualization layout—called the *integrated ecosystem* layout—therefore uses a force-directed visualization approach using OpenOrd to create clusters and Force Atlas to aesthetically space nodes [Bastian et al. 2009]. A pseudocode of our integrated ecosystem layout is provided in Algorithm 2. We used the same node and edge color coding and node and edge sizing for both visualizations. Node and relationship types are differentiated by color (e.g., red = firms; green = investments and investment firms; blue = people). Nodes are sized by betweenness centrality and edge thickness corresponds to the number of internode relationships. The pseudocode for these helper functions is shown in Algorithm 3.

3.2.4. Step 4: Sense Making and Storytelling. The true purpose of data-driven visualization is more than pretty pictures [Card et al. 1999] and does not eliminate the need for human insight and foresight [McAfee and Brynjolfsson 2012]. While visualization is primarily about data transformation, representation, and interaction, ultimately it is about harnessing human visual perception capabilities to help identify trends, patterns, and outliers with computational capabilities [Card et al. 1999]. It involves the formation of abstract visual metaphors in combination with a human information discourse (interaction) that enables detection of the expected and discovery of the

ALGORITHM 1: Generating *bicentric ecosystem* layout

input : 2-step egonetworks of two focal entities and relationships among all nodes in the egonetworks
output : Node position, node size, node color, edge weight, and edge color for the bicentric layout anchoring around two focal nodes

$G, n_1, n_2 \leftarrow$ the entire graph containing all nodes and edges, focal node 1, focal node 2;
$M_1, M_2 \leftarrow$ constants;
// **Determine node size, node color, edge weight, and edge color**;
foreach *node* $\in$ *allNodes* **do**
  nodeSize[node] $\leftarrow k$ `Betweenness`$(G, node)^{\alpha}$;
  nodeColor[node] $\leftarrow$ `ComputeNodeColor`(*node*);
end
foreach *edge* $\in$ *allEdges* **do**
  $node_1, node_2 \leftarrow$ Two terminal nodes of *edge*;
  edgeWeight[edge] $\leftarrow$ `EdgeDegree`$(node_1, node_2)$;
  edgeColor[edge] $\leftarrow$ `ComputeEdgeColor`(*edge*);
end
// **Determine for each node how many steps away from the focal nodes**;
foreach *node* $\in$ *allNodes* **do**
  steps[node] $\leftarrow (0, 0)$;
  if *node is one step away from* n_1 **then** $steps[node][1] \leftarrow 1$;
  else if *node is two step away from* n_1 **then** $steps[node][1] \leftarrow 2$;
  if *node is one step away from* n_2 **then** $steps[node][2] \leftarrow 1$;
  else if *node is two step away from* n_2 **then** $steps[node][2] \leftarrow 2$;
end
// **Extract largest connected components, determine cluster size, and compute node type composition**;
foreach *(p,q)* $\in \{(2,2),(1,2),(2,1)\}$ **do**
  $G_{p,q} \leftarrow \{node \mid node \in G \wedge steps[node] = (p, q)\} \cup \{edge \mid steps[node_1] = (p, q) \wedge steps[node_2] = (p, q)\}$;
  $topNodes_{p,q} \leftarrow$ nodes in the largest connected component of $G_{p,q}$;
end
foreach *(p,q)* $\in \{(1,1),(2,2),(1,2),(2,1)\}$ **do**
  If $(p, q) = (1, 1)\beta \leftarrow \sqrt{2}\beta \leftarrow 1$;
  $clusterSize_{p,q} \leftarrow \beta M_1 \sqrt{\sum_{\forall node \in G_{p,q}} \texttt{Degree}(G, node)}$;
end
foreach *(p,q)* $\in \{(1,0),(2,0),(0,1),(0,2)\}$ **do**
  $R_{p,q}[1] \leftarrow \sum_{node \in G_{p,q} \cap type = \text{Firm}} \texttt{Degree}(G, node)$;
  $R_{p,q}[2] \leftarrow \sum_{node \in G_{p,q} \cap type = \text{Investor}} \texttt{Degree}(G, node)$;
  $R_{p,q}[3] \leftarrow \sum_{node \in G_{p,q} \cap type = \text{Person}} \texttt{Degree}(G, node)$;
  $R_{p,q} \leftarrow R_{p,q} / \sum R_{p,q}$;
end
// **Compute node positions**;
foreach *node* $\in$ *allNodes* **do**
  if $node = n_1$ **then** $nodePos[node] \leftarrow (-1, 0)$;
  else if $node = n_2$ **then** $nodePos[node] \leftarrow (1, 0)$;
  else if $steps[node] \in \{(1, 0), (2, 0), (0, 1), (0, 2)\}$ **then**
    $p, q \leftarrow steps[node][1], steps[node][2]$;
    switch $type[node]$ **do**
      case *Firm* $\theta \leftarrow \pi \texttt{Random}(\pi) R_{p,q}[1] - \pi/2$;
      case *Investor* $\theta \leftarrow \pi(R_{p,q}[1] + \texttt{Random}(\pi) R_{p,q}[2]) - \pi/2$;
      case *People* $\theta \leftarrow \pi(R_{p,q}[1] + R_{p,q}[2] + \texttt{Random}(\pi) R_{p,q}[3]) - \pi/2$;
    endsw
    $nodePos[node] \leftarrow (1 + \cos\theta \times (p + q), \sin\theta \times (p + q))$;
    if $steps[node][2] = 0$ **then** $nodePos[node] \leftarrow -nodePos[node]$;
    else $nodePos[node][2] \leftarrow -nodePos[node][2]$;
  else if $steps[node] = (1, 1)$ **then** $nodePos[node] \leftarrow \texttt{Jitter}((0, 0), clusterSize_{1,1})$;
  else if $steps[node] = (2, 2)$ **then**
    $nodePos[node] \leftarrow \texttt{Jitter}((0, \sqrt{3}), clusterSize_{2,2})$;
    if $node \notin topNodes_{2,2}$ **then** $nodePos[node][2] \leftarrow -nodePos[node][2]$;
  else if $steps[node] = (1, 2)$ **then**
    $nodePos[node] \leftarrow \texttt{Jitter}((-3/4, \sqrt{15}/4), clusterSize_{1,2})$;
    if $node \notin topNodes_{1,2}$ **then** $nodePos[node][2] \leftarrow -nodePos[node][2]$;
  else if $steps[node] = (2, 1)$ **then**
    $nodePos[node] \leftarrow \texttt{Jitter}((3/4, \sqrt{15}/4), clusterSize_{2,1})$;
    if $node \notin topNodes_{2,1}$ **then** $nodePos[node][2] \leftarrow -nodePos[node][2]$;
end

ALGORITHM 2: Generating *integrated ecosystem* layout

```
input  : Aggregated 2-step egonetworks of all focal entities and relationships among all
         nodes
output : Node position, node size, node color, edge weight, and edge color for the full
         network layout
G ← the entire graph containing all nodes and edges;
// Determine node size, node color, edge weight, and edge color;
foreach node ∈ allNodes do
    nodeSize[node] ← k√Degree(G,node);
    nodeColor[node] ← ComputeNodeColor(node);
end
foreach edge ∈ allEdges do
    node1, node2 ← Two terminal nodes of edge;
    edgeWeight[edge] ← EdgeDegree(node1,node2);
    edgeColor[edge] ← ComputeEdgeColor(edge);
end
// Determine initial node positions using OpenORD layout algorithm;
nodePos ← OpenORD(G);
// Jitter nodes to reduce node overlaps using Noverlap layout algorithm;
nodePos ← Noverlap(nodePos,nodeSize);
```

unexpected within massive, dynamically changing information spaces [Wong and Thomas 2004].

Sense making has its roots in cognitive psychology and many different models have been developed [North 2006]. The consensus across these models is that the sense-making process is cyclic and interactive, involving both discovery and creation. During the generation loop, an individual searches for representations. In the data coverage loop, we instantiate these representations. Based in these insights, we shift our representation and begin again. Following data/frame theory [Klein et al. 2006], the sense-making process takes place within a frame in our study the business ecosystem. The “frame” manages our attention and we define, connect, and filter the data. Visualizations facilitate this process. By tracking anomalies, detecting inconsistencies, judging plausibility, and gauging data quality, we question the frame. If we disagree with the frame, we go back to the data and reframe. If we agree with the frame, we preserve it and elaborate on it by adding and filling slots, seeking inferring data, discovering new data and relationships, and discarding data. Through iteration and discussion, stakeholders may detect a pattern and match and/or adjust their frame or mental model. Together this forms a complete sense-making loop.

Sense making can more generally be described as the process from data to understanding, taking into account both the fact that a variety of senses may legitimately emerge from the same data and that closing the gap between data and theory may include iterations [Langley 1999]. Overall, the different strategies for sense making provide a useful repertoire of approaches for attacking the vast amount of data that is generally generated through analysis [Bizzi and Langley 2012]. Our visualizations of metrics and networks can therefore be seen to support the sense-making process by providing interactive representations for the complex concept of an ecosystem. By providing a means to model the skeleton of an ecosystem, they contribute to a collectively realized shared vision [Hagel and Brown 2005] and serve to build a shared vision of antecedent or subsequent events [Russell et al. 2011].

ALGORITHM 3: Helper functions

```
// Basic functions;
Function(Random(maxRange))
return maxRange × uniform random number ∈ [0, 1);
Function(Jitter((centerX,centerY),radius))
φ, r ← Random(2π), radius√Random(1);
return (centerX+r cos φ, centerY+r sin φ);
// Graph-theoretic measure computation;
FunctionDegree(graph,node)
return weighted degree of node in graph;
FunctionEdgeDegree(node1,node2)
return number of relationships between node1 and node2;
FunctionBetweenness(graph,node)
return betweenness centrality of node in graph;
// Existing graph layout algorithms;
FunctionNoverlap(currentNodePos,currentNodeSize)
return nodePosByNoverlap;
FunctionOpenORD(graph)
return nodePosByOpenORD;
// Node and edge color assignment;
Function(ComputeNodeColor(node))
switch type[node] do
    case Firm return Red;
    case Investor return Green;
    case People return Blue
endsw
FunctionComputeEdgeColor(edge)
switch type[edge] do
    case Acquisition return Red;
    case Investment return Green;
    case Contribution return Blue
endsw
```

Visualizations are highly contextual; they are interpreted in the context of the user with user actions such as inspecting, ranking, comparing, categorizing, inferring, associating, and correlating [Xu et al. 2009], and they are influenced by the various inherent biases of each individual, as well as by the groups that participate in sense-making activities. To support the sense making and subsequent storytelling, or visual narratives intended to convey stories [Segel and Heer 2010], fundamentals of telling a good story (such as making sure the audience has enough background knowledge about the specific dataset, showing the visual elements and what they represent, arranging in a way that tells an interesting story) have been found to lessen the difficulties arising from introducing visualizations to the broader public [Ma et al. 2012].

4. ILLUSTRATIVE EXAMPLES

We illustrate our data-driven visualization methodology for understanding business ecosystem dynamics with three different relational contexts—a corporate acquisition, firm coopetition, and a strategic alliance—in the mobile industry. Specifically, we focus on the acquisition of Motorola Mobility by Google in 2011, the coopetition between Apple and Samsung (2010–2012), and the strategic alliance formed between Nokia

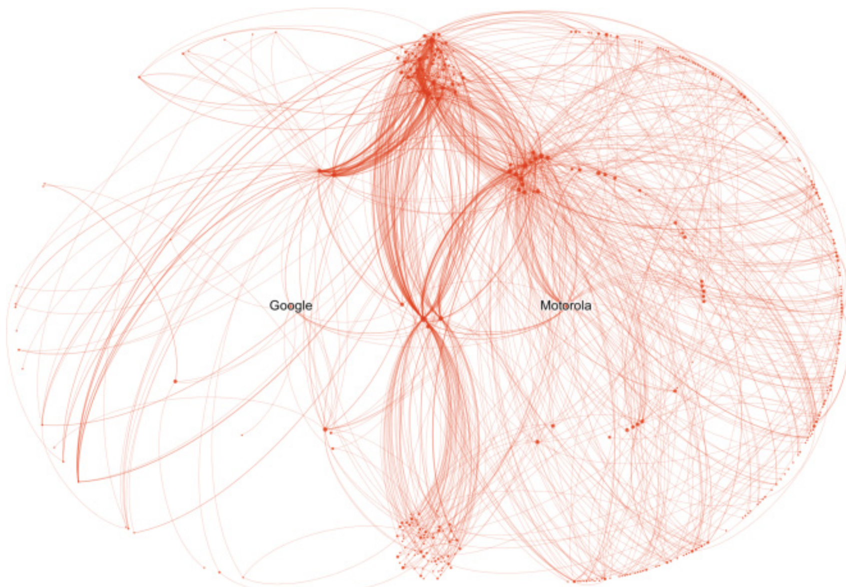

Fig. 3. Bicentric DAR ecosystem of Google and Motorola Mobility.

and Microsoft in 2011. We purposely selected these six focal companies for our study as they are considered key players in the mobile ecosystem and are shaping the growth and trajectory of the industry. They exhibit similar organizational characteristics in terms of size, employees, and scope. Each of the players entered the mobile ecosystem at different times, and their ecosystems suggest they entered with different objectives, constraints, and strategies. We examine the formation and evolution of the three cases over a 13-year period (2000–2012) as it marks a period of the birth, rapid growth, and hypercompetition in the mobile industry. For each focal firm, we identify the entire two-step network.

4.1. Acquisition: Google and Motorola Mobility

Google's acquisition of Motorola Mobility in August 2011 received significant attention from players in the mobile ecosystem. Motorola Mobility had been struggling to (re)gain market share in the lucrative smartphone segment. Through various enterprise transformations over the years, it had tried to reposition itself but still failed to deliver on its past innovative pedigree. On the contrary, Google—not a traditional mobile player—was speculated on many occasions to be entering the mobile ecosystem in full force. For instance, Google had been a key bidder on the wireless spectrum in 2008. More recently, Google was a key investor in and creator of the Android mobile platform. However, there were no signs that Google would offer its own hardware.

Figures 3 and 4 show the bicentric ecosystem layout of Google and Motorola Mobility using DAR and EFR data, respectively. In the DAR visualization, we notice that the two focal firms share only a few direct partners. Motorola Mobility's two-step network is much denser than that of Google. Interestingly, we observe that two key clusters—Google and Motorola Mobility—share a significant number of second-tier ecosystem partners who themselves are highly interconnected (i.e., are in the main component). Similarly, we see that many of Motorola's first-tier partners are in fact in Google's second tier. In the EFR visualization, the patterns are virtually reversed. Google's two-step network is vastly more dense and the two focal firms share no direct and only

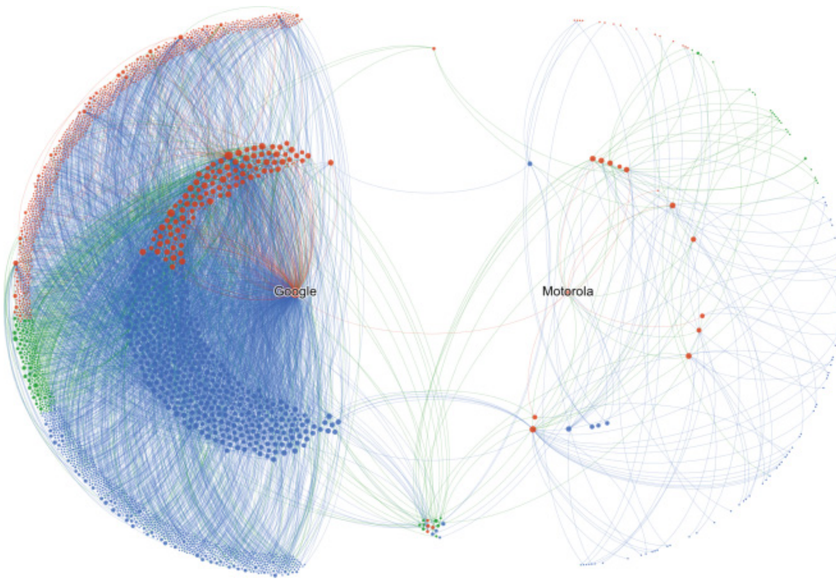

Fig. 4. Bicentric EFR ecosystem of Google and Motorola Mobility.

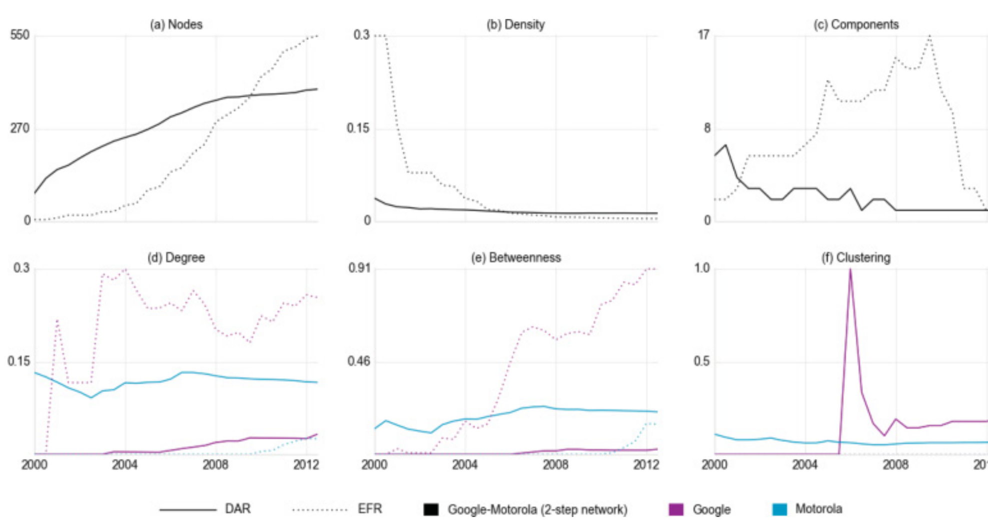

Fig. 5. Google and Motorola Mobility: small multiples of bicentric ecosystem metrics.

a small number of indirect partners (people, firms, and investors). We also note that Google's second-tier network is rather large with a substantial number of investment firms (depicted in green).

Figure 5 reveals the longitudinal evolution of this ecosystem. At the network level (a–c), we observe that the density of deals and alliance relationships is stable; nodes increase and then stabilize after 2008, when components also stabilize. The density of executive and funding relationships decreases precipitously at the beginning of the decade; the increase in nodes accelerates after 2007; and components increase steadily until they drop dramatically after 2010. Among the entities with which Google and Motorola (and their direct connections) had relationships during the 2000–2012 period, the number of entities continued to grow even though a maintainable equilibrium had been reached. A serious reduction in components accompanies the acquisition of Motorola Mobility by Google in the 2010–2011 period.

At the node level (d–f), the EFR activity of Google contrasts with the EFR stability of Motorola in their respective networks. In the agile Silicon Valley context, fluctuating connections to Google are shown in the EFR degree values over the 2000–2012 period, with concomitant increases in betweenness values as Google's aristocratic position becomes dominant. The 2005 spike in Google's clustering coefficient, when its acquired companies had connections to each other, is followed by a pattern of acquisitions with more modest but still significant interrelated deal and alliance relationships. This conforms with the pattern shown in the bicentric network visualizations earlier.

An analysis of the POD data shown in Table III reveals that a high proportion of the topics mentioned involves the Android operating system, smartphones, lawsuits and litigation, patent protection patent litigation, and intellectual property. These topics underline that Google's acquisition of Motorola Mobility were primarily associated with gaining traction in the smartphone industry. The particularly high comention of Apple, Samsung, and Nokia also reveals the event's relevance to other key handset vendors.

The triangulation of DAR, EFR, and POD data reveals three different perspectives on the Motorola/Google ecosystem. A few direct deals and alliances existed (DAR), most executive and investment relationships (EFR) were indirect, and legal issues dominated the public attention to the two companies (POD).

4.2. Coopetition: Apple and Samsung

Apple and Samsung both serve the consumer electronics market. Apple's business extends to computer hardware and software, as well as digital distribution. Samsung's business extends to home appliances, telecom equipment, and semiconductors. In the mobile ecosystem, Samsung was a major component provider of Apple's products until recently. Samsung's smartphones use the competing Android platform, and Samsung is arguably the fiercest competitor of Apple in the mobile device segment, highlighted by an ongoing intense legal battle between the two firms.

Figures 6 and 7 show the bicentric ecosystem layout of Apple and Samsung using DAR and EFR data, respectively. In the DAR visualization, we observe several interesting patterns. Apple and Samsung share several direct partners as well as many indirect (first- and second-tier) partners. This points to the highly interconnected and complex nature of their relationship. While Apple and Samsung were partners, they are also competitors. At the same time, they have a significant amount of overlap in their two-step network. In the EFR visualization, we notice that the two firms do not share any direct partners and only a few indirect ones. Apple has a much larger EFR network than Samsung, potentially pointing to its proximity to Silicon Valley and different governance structure.

Figure 8 provides the evolution of metrics for this focal firm dyad ecosystem. At the network level (a–c), we observe a steady state in the density of deal and alliance relationships, along with a gradual but tapering increase in the number of nodes and a stabilization of components following 2008. A more modest increase is shown in the number of executive and funding relationships (nodes) connected directly and indirectly to Apple and Samsung during this period. A precipitous drop in the density of relationships following 2001 (following the 2001 digital bust) and a dramatic increase in the components are evidenced, culminating in 2006 to 2009 (which corresponds to Apple's aggressive growth in content distribution).

At the node level (d–f), Apple's proliferation of executive and investor relationships between 2000 and 2004 is reflected in its density and betweenness metrics for Apple. These are in contrast to the steady levels of these relationships for Samsung and the

Table III. Google and Motorola Mobility: Mentions of Business Issues, Companies, Venture-Funded Companies, and Technologies (Count in Brackets) 2010–2012

	Business Issue	Companies	Venture Funded Companies	Technologies
1	Lawsuits and Litigation [9,115]	Google [37,561]	Xoom [2,723]	Android Operating System [26,348]
2	Acquisitions [6,380]	Motorola [37,412]	Twitter [1,166]	Smartphones [21,190]
3	Patent Protection [4,891]	Apple [21,982]	Dropbox [184]	Tablet Computing [16,363]
4	Patent Litigation [4,867]	Samsung [17,433]	ViewSonic [163]	Fourth-Generation Wireless (4G) [5,015]
5	Intellectual Property [4,039]	Microsoft [10,099]	Foursquare [145]	Ice Cream Sandwich (Android) [4,326]
6	Ecosystem [3,872]	Nokia [6,507]	Reddit [115]	Mobile Computing [4,039]
7	Technology Leader [2,892]	Verizon [5,581]	Evernote [102]	Electronic Books and Devices [3,196]
8	International Trade [2,869]	Amazon.com [4,699]	Jawbone [96]	Third-Generation Wireless (3G) [2,832]
9	New Products [2,585]	LG [4,558]	Roku [95]	Social Networking [2,573]
10	Rumors [2,501]	Research In Motion [4,398]	Rovio [94]	Mobile Phones [2,339]
11	Antitrust Action [2,290]	AT&T [3,684]	Pinterest [83]	Microsoft Windows 8 [2,294]
12	Merger [1,956]	Hewlett-Packard [3,312]	Flipboard [76]	Jelly Bean (Android) [2,135]
13	Strategic Partnerships [1,859]	Intel [2,777]	Tumblr [73]	Dual-Core Processors [1,840]
14	Market Leader [1,501]	Sony [2,501]	OnLive [72]	Touchscreen [1,818]
15	Shift in Market Share [1,197]	Facebook [2,475]	Flurry [70]	Cloud Computing [1,592]
16	Strategic Planning [1,045]	Sprint Nextel [2,299]	Bit9 [57]	Set-Top Box [1,480]
17	Business Model [913]	T-Mobile [2,200]	Sonos [57]	Open Source Software [1,426]
18	Staff Reduction [805]	Nortel [1,977]	Shazam [51]	Microsoft Phone 7 [1,313]
19	Benchmarks [633]	Oracle [1,828]	Rdio [50]	Voice Assistant [1,174]
20	Market Size and Growth [613]	Qualcomm [1,729]	Appcelerator [49]	Personal Computers [1,161]

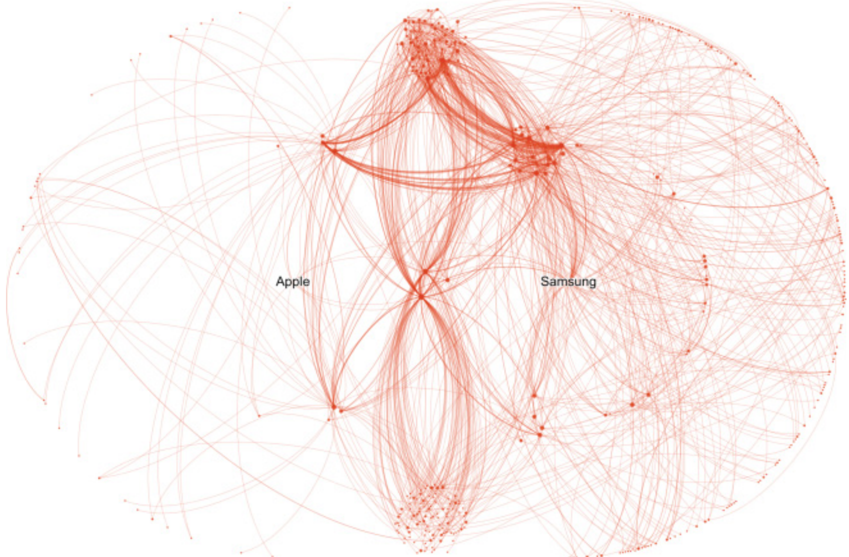

Fig. 6. Apple and Samsung bicentric DAR ecosystem.

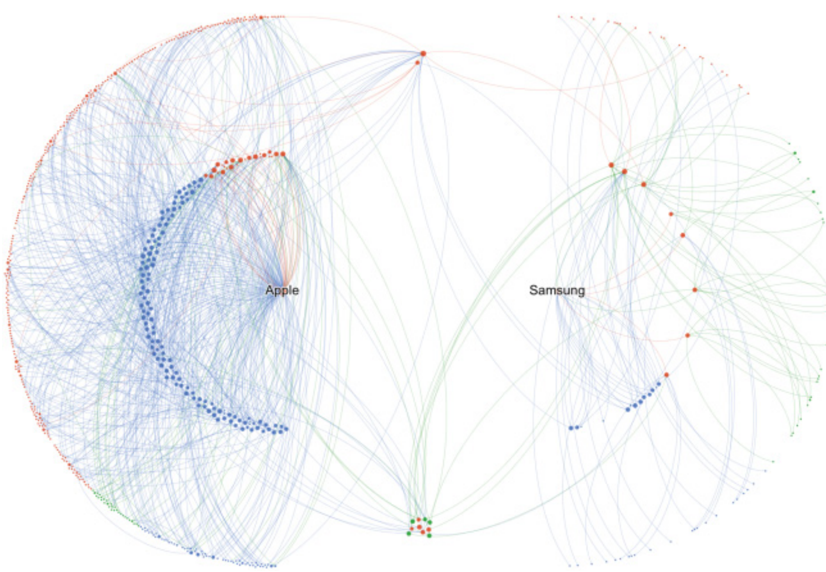

Fig. 7. Apple and Samsung bicentric EFR ecosystem.

relatively steady levels of deal and alliance relationships for both Apple and Samsung during this period. The EFR betweenness values for both Apple and Samsung increase following 2008, at a time when Samsung's CEO changed. The fluctuation in clustering coefficients for Samsung settles into a steady state following 2004, at which time Apple's clustering values spike, level off, then spike again in 2011 (when Apple's CEO changed).

The Apple/Samsung ecosystem is seen differently through the lenses of DAR, EFR, and POD data, yet each reveals a unique aspect of the changes in Apple and Samsung's relationships during the 2000–2012 period, illustrating the responsiveness to change

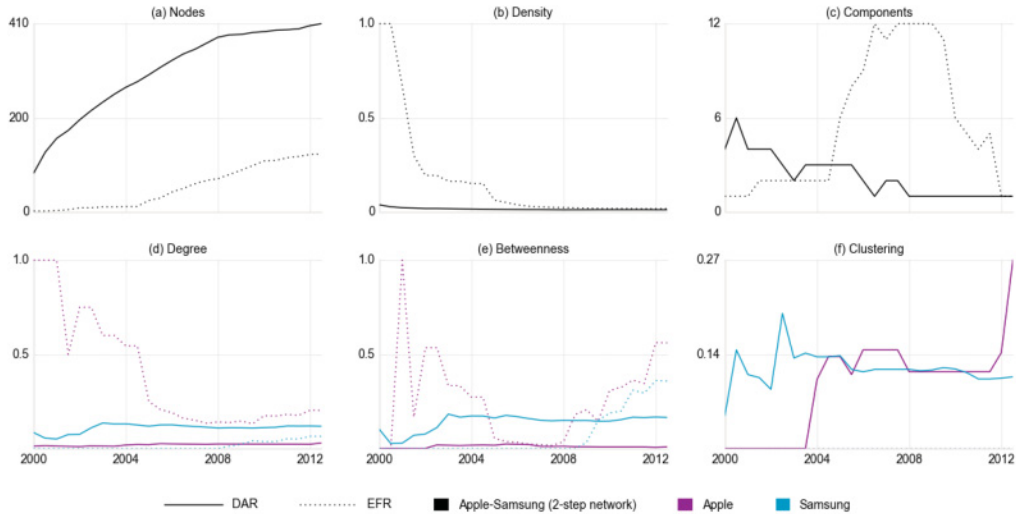

Fig. 8. Apple and Samsung bicentric ecosystem metrics.

that characterizes their relationship of coopetition. The cooperation of Apple and Samsung as customer/supplier in the early part of the decade underwent several periods of significant change as market conditions sent tremors through the investment communities of Apple during 2001 and 2008, during 2009 when Samsung's CEO changed, and in 2011 when Apple's CEO changed. Apple and Samsung both held significant (but complementary) aristocratic positions in different ecosystems. They responded differently to the same external forces in the larger ecosystem. Apple's responsiveness to the investment community was to tighten reigns (reducing degree and increasing clustering), and Samsung's response to its broad-based vendors and supplier relationships was to hold a steady course. Legal issues dominate the press mentions (POD) of Apple and Samsung between 2010 and 2012 (Table IV), and the metrics of 2012 in Figure 8 hint that Apple may pursue the course of further tightening.

4.3. Strategic Alliance: Nokia and Microsoft

The alliance between Nokia and Microsoft in February 2011 was considered by many pundits to be an inevitable move given the recent struggles of both companies in the mobile ecosystem. Once a leader in the global handset market, Nokia began in 2007 to fall behind other device manufacturers in the lucrative smartphone segment. Microsoft, a perennial leader in the desktop market, had achieved limited traction in the mobile market despite its Windows Mobile platform. Many attributed Microsoft's shortcoming to a lack of an appropriate hardware partner. Though with high initial hopes, the success of the strategic alliance has been questioned repeatedly, especially as market share of Nokia Windows phones has remained low and Nokia's share price has plummeted.

Figures 9 and 10 show the bicentric ecosystem layout of Nokia and Microsoft using DAR and EFR data, respectively. In the DAR visualization, we observe the highly interconnected nature of the two companies. Not only do the firms have direct relationships, but also they share many direct and indirect partners. In fact, the visualization reveals that the majority of Nokia's DAR network is in fact subsumed in Microsoft's two-step network. In the EFR visualization, we notice that several individuals connect the two companies, including Steven Elop. Microsoft's EFR network is vastly richer and denser than that of Nokia.

Table IV. Apple and Samsung: Mentions of Business Issues, Companies, Venture-Funded Companies, and Technologies (Count in Brackets) 2010–2012

	Business Issue	Companies	Venture Funded Companies	Technologies
1	Lawsuits and Litigation [25,195]	Samsung [99,978]	Twitter [2,372]	Smartphones [59,343]
2	Patent Litigation [12,301]	Apple [98,059]	Xoom [1,841]	Tablet Computing [54,020]
3	New Products [8,928]	Google [38,581]	Dropbox [626]	Android Operating System [42,222]
4	Intellectual Property [7,462]	Nokia [20,878]	Roku [527]	Fourth-Generation Wireless (4G) [11,167]
5	Rumors [7,218]	Motorola [19,283]	Rovio [282]	Electronic Books and Devices [9,970]
6	Ecosystem [6,469]	Microsoft [18,131]	Evernote [268]	Microsoft Windows 8 [8,723]
7	Technology Leader [6,467]	Amazon.com [13,929]	ViewSonic [265]	Mobile Computing [8,315]
8	Patent Protection [4,659]	LG [11,953]	Foursquare [219]	Third-Generation Wireless (3G) [7,806]
9	Market Leader [4,173]	Research In Motion [10,263]	Reddit [202]	Mobile Phones [6,249]
10	International Trade [3,946]	Sony [9,982]	Pinterest [191]	Ice Cream Sandwich (Android) [5,024]
11	Acquisitions [3,617]	AT&T [9,255]	Flipboard [189]	Touchscreen [4,596]
12	Strategic Partnerships [3,598]	Verizon [8,611]	Flurry [187]	Voice Assistant [4,439]
13	Shift in Market Share [3,402]	Hewlett-Packard [7,357]	OnLive [185]	Quad-Core Processors [4,371]
14	Market Size and Growth [3,038]	Intel [7,151]	Tumblr [158]	Personal Computers [3,908]
15	Benchmarks [2,116]	T-Mobile [5,136]	Boxee [157]	Dual-Core Processors [3,766]
16	Price Cuts [2,033]	Sprint Nextel [5,092]	Sonos [147]	Microsoft Phone 7 [3,443]
17	Antitrust Action [1,910]	Qualcomm [4,707]	Jawbone [141]	Social Networking [3,289]
18	Strategic Planning [1,766]	Dell [4,479]	SpaceX [134]	Cloud Computing [2,976]
19	Business Model [1,713]	Lenovo [4,196]	Shazam [108]	Jelly Bean (Android) [2,720]
20	Staff Reduction [1,666]	HTC [3,989]	Appcelerator [105]	High-Definition Television (HDTV) [2,412]

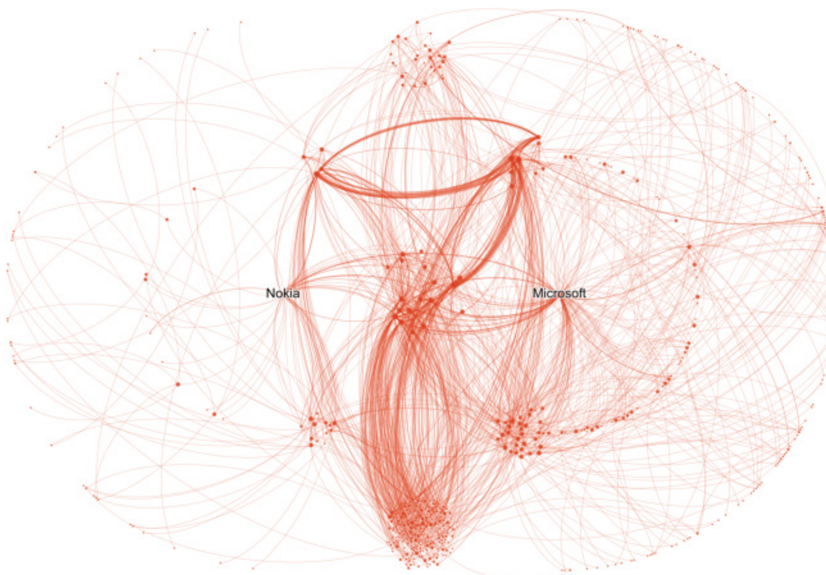

Fig. 9. Nokia and Microsoft bicentric DAR ecosystem.

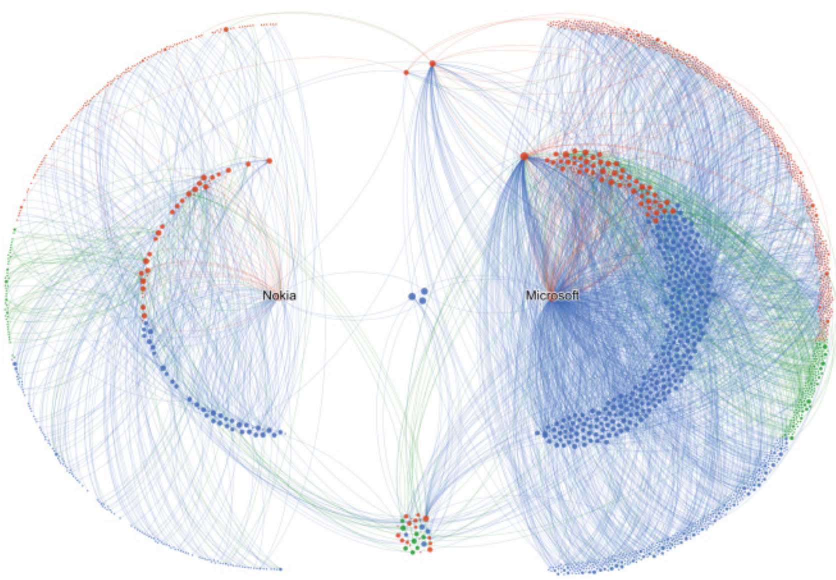

Fig. 10. Nokia and Microsoft bicentric EFR ecosystem.

Figure 11 reveals the longitudinal evolution of this ecosystem. At the network level (a–c), the DAR and EFR metrics show similar patterns in the increasing number of nodes and the decreasing density. The components metric for EFR shows dramatic increases during the 2004–2007 period, similar to the other two cases. Microsoft's and Nokia's individual two-step networks (d–f) show patterns of synchrony, even though the levels are different. In both DAR and EFR relationships, the degree values fluctuate moderately and stabilize in 2008. The betweenness values for DAR of the two focal

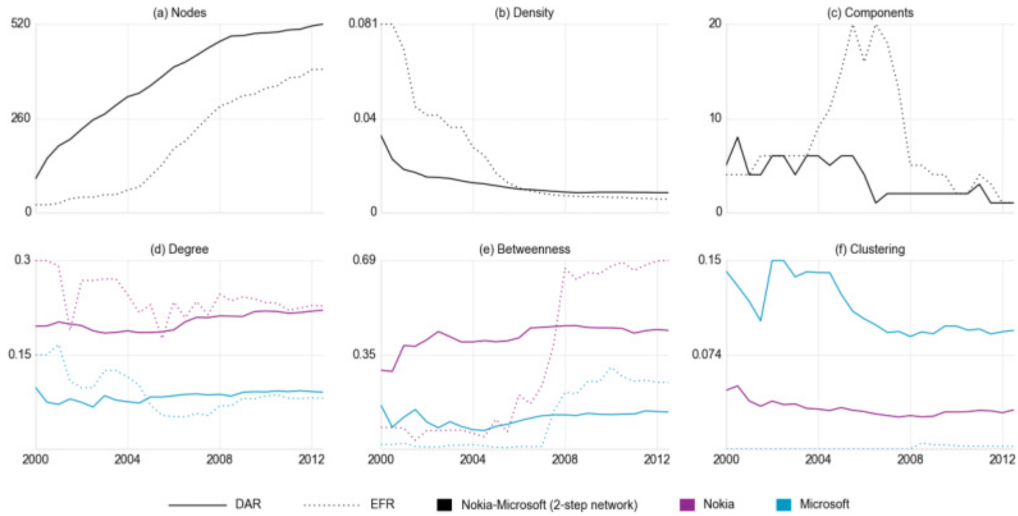

Fig. 11. Nokia and Microsoft bicentric ecosystem metrics.

firms are relatively stable, although at different levels, while the betweenness values for EFR increase dramatically in 2008. Clustering coefficients for EFR remain steady over this period for both. In the 2000–2005 period, the DAR clustering coefficients for Nokia show more fluctuation than those for Microsoft, and the clustering coefficients for Nokia, although relatively steady from 2006 through 2010, remain higher than those for Microsoft. These findings correspond to the network visualizations shown in Figures 9 and 10.

The triangulation of DAR, EFR, and POD data reveals three different yet complementary perspectives of the Microsoft/Nokia ecosystem. Seen from a network perspective, the alliance between Microsoft and Nokia was based on a fabric of relationships that reflected synchrony in their respective (and shared) ecosystems, even in the face of the external market disruptions in 2001 and 2008. The DAR and EFR network perspectives reveal similar patterns of relationships between the Microsoft and Nokia relationships that could reflect shared understanding of current and future conditions and of the mutual business opportunities. The public discourse, revealed in POD, supports this, with ecosystem and strategic partnerships outpacing mentions of lawsuits and litigation, as shown in Table V.

4.4. Summary

The companies, individuals, and investors shown in the networks of three examples are embedded in a large complex system of relationships. Changes in the network over time reveal what Laughlin and Pines [2000] refer to in physics as complex adaptive behavior as the emergent structure evolves. The complexity is dynamic and self-organizing in ways that heavily influence the probabilities of later events [Prigogine 1997].

The dynamic, nonlinear, and complex mobile ecosystem shows aristocratic patterns, in which powerful individuals and organizations play a pivotal role in the way the ecosystem evolves over time [Barabási 2003] until they trigger an avalanche of changes, a tipping point [Gladwell 2000]. Figures 12 and 13 show the integrated ecosystem of all six focal companies using both DAR and EFR in 2004 and 2012, respectively. The two time snapshots reveal the significant changes that occurred over time. Microsoft's dominant role in the many deals and alliances that characterize its network reveals its aristocratic position. Google's high degree of centrality with links to key individuals,

Table V. Nokia and Microsoft: Mentions of Business Issues, Companies, Venture-Funded Companies, and Technologies (Count in Brackets) 2010–2012

	Business Issue	Companies	Venture Funded Companies	Technologies
1	Ecosystem [4,718]	Nokia [34,943]	Twitter [1,031]	Smartphones [23,084]
2	Strategic Partnerships [4,087]	Microsoft [23,974]	Rovio [263]	Android Operating System [16,730]
3	New Products [3,258]	Apple [19,592]	Dropbox [247]	Tablet Computing [12,050]
4	Lawsuits and Litigation [3,244]	Google [15,171]	Foursquare [218]	Microsoft Windows 8 [7,560]
5	Rumors [2,722]	Samsung [14,084]	Xoom [196]	Microsoft Phone 7 [6,497]
6	Technology Leader [2,704]	AT&T [6,211]	Appcelerator [127]	Mobile Computing [3,687]
7	Market Leader [2,127]	Research In Motion [5,879]	Waze [115]	Fourth-Generation Wireless (4G) [3,233]
8	Acquisitions [1,759]	Motorola [5,771]	Cyan [113]	Mobile Phones [3,007]
9	Intellectual Property [1,597]	Amazon.com [3,615]	ViewSonic [105]	Electronic Books and Devices [2,088]
10	Patent Protection [1,543]	Verizon [3,500]	Evernote [95]	Social Networking [1,949]
11	Shift in Market Share [1,489]	Intel [3,166]	Reddit [95]	Cloud Computing [1,857]
12	Patent Litigation [1,371]	LG [2,836]	Pinterest [74]	Third-Generation Wireless (3G) [1,804]
13	Strategic Planning [1,361]	T-Mobile [2,763]	Shazam [69]	Microsoft Windows 7 [1,491]
14	International Trade [1,232]	Hewlett-Packard [2,729]	YCharts [56]	Linux [1,490]
15	Market Size and Growth [1,020]	Qualcomm [2,480]	OnLive [53]	Touchscreen [1,452]
16	Staff Reduction [948]	Sony [2,059]	LivingSocial [51]	Dual-Core Processors [1,398]
17	Business Model [803]	Facebook [1,927]	SpaceX [51]	Personal Computers [1,284]
18	General Cost Reduction [773]	IBM [1,884]	Flipboard [50]	HTML5 (Programming Language) [1,163]
19	Regulatory Action [764]	Huawei [1,757]	Flurry [49]	Blackberry 10 [1,162]
20	Price Cuts [675]	Sprint Nextel [1,570]	Jawbone [46]	Open Source Software [1,028]

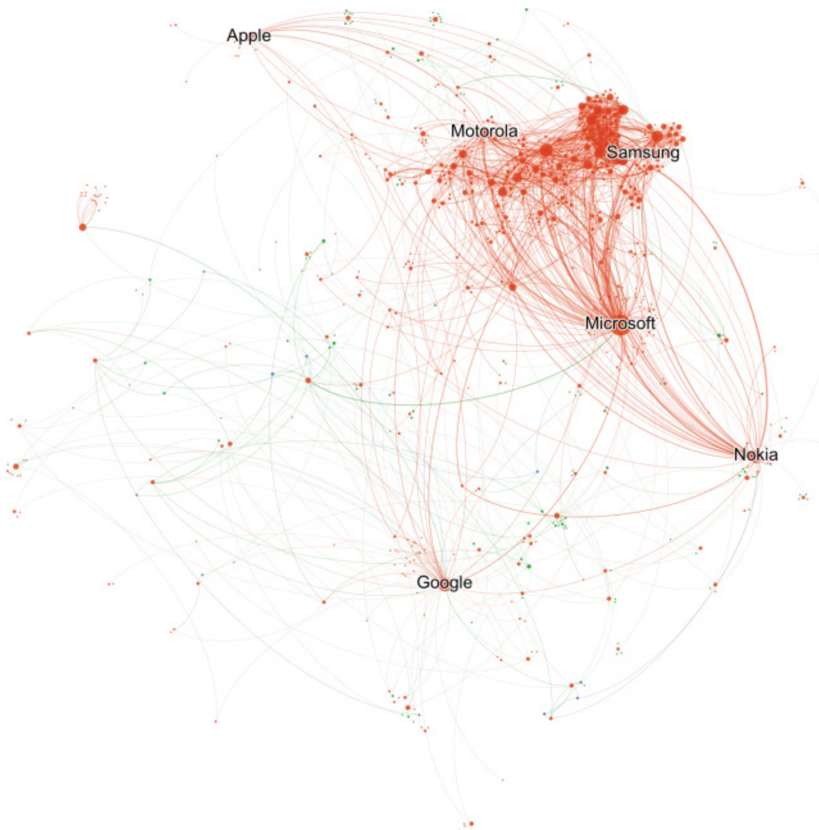

Fig. 12. Mobile ecosystem (2004).

many of whom became Googlers as their startups were acquired by Google, illustrates the dominant role of a relationship hub, a node with a very large number of links. The Google hub illustrates the rich-get-richer condition in which the pivotal role of the hub carries a disproportionate influence [Buchanan 2002].

Against this context, the patterns revealed by Apple and by Samsung show well-established positions in the network with many structural holes. An absence of structural holes [Burt 2009], illustrated by the edges between Microsoft and Nokia, shows a cluster of nodes interlinked in a pattern that allows easy communication and coordination among the node's links. They create a powerful negotiating position, in contrast to a node with many structural holes, such as the Google/Motorola pattern in which unconnected nodes can play against each other, dividing and conquering, giving a strong rationale for Google's acquisition of Motorola to preclude either of its competing platforms Microsoft or Apple from acquiring access to Motorola's unique technologies.

In the context of the whole ecosystem, a node's power becomes a function of all related nodes' power, including those far away from the node [Borgatti et al. 2009], such as the seemingly disparate networks of Google and Motorola. Interdependent systems structure the nature of a complex business ecosystem through multiple codependencies (moorings, if you will) that enable the fluidities to be realized without chaos [Urry 2004]. As a self-organizing system, entities in an ecosystem are coevolving and interdependent.

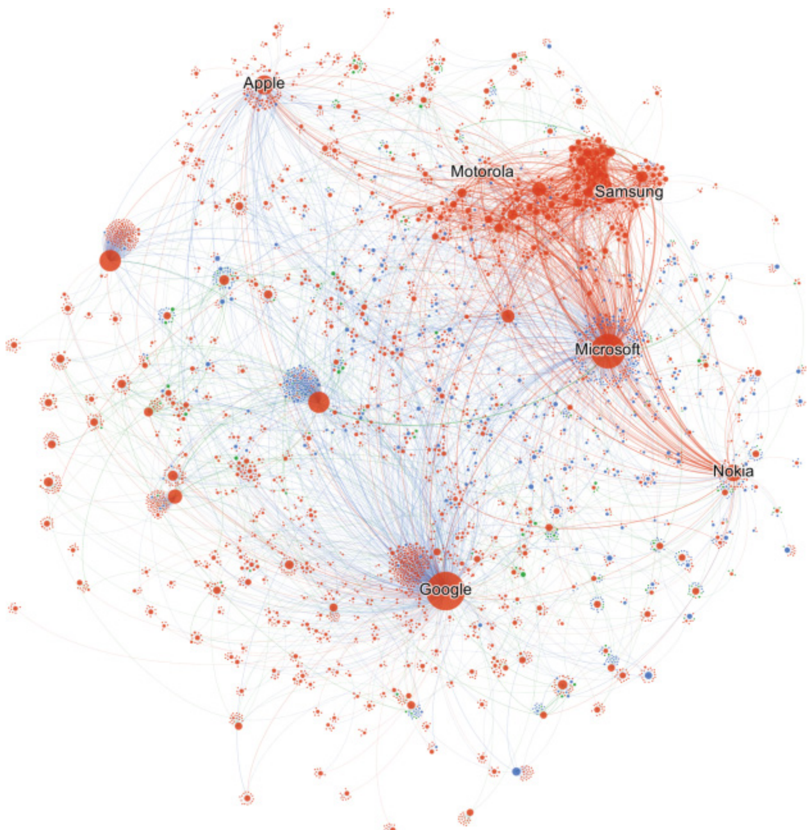

Fig. 13. Mobile ecosystem (2012).

Figure 14 shows the common nodes found in the DAR and EFR datasets. It reveals that only a very small fraction of firms for this ecosystem are present in both datasets, but that together it can reveal a much broader picture.

5. DISCUSSION

This study has introduced a methodology with which to analyze and visualize the structure and dynamics of business ecosystems. We demonstrated our approach using examples from the mobile industry using three sources of data that reflect different realities of evolving business ecosystems. Data about deals and alliances (DAR) were used to reflect the institutional perspective; completed agreements and deals between companies in the institutionally curated SDC dataset were gleaned from methodically documented sources. Data about relationships at the executive and funding levels of companies (EFR) were used from the IEN dataset to reflect the perspective of personal relationship capital. These institutionally curated and socially constructed datasets revealed both formal and informal ties and offered insights about the fluidity of the ecosystem—the flow of transactions, information, talent, and financial resources. Data from Northern Lights records of press mentions of companies in the three selected cases reflected the public opinion and discourse (POD) of the mobile ecosystem.

There are many challenges and opportunities for data-driven analysis and visualization of business ecosystems. Arguably, the most foundational task is the careful selection and orchestrated use of multiple datasets. Datasets often use different unique

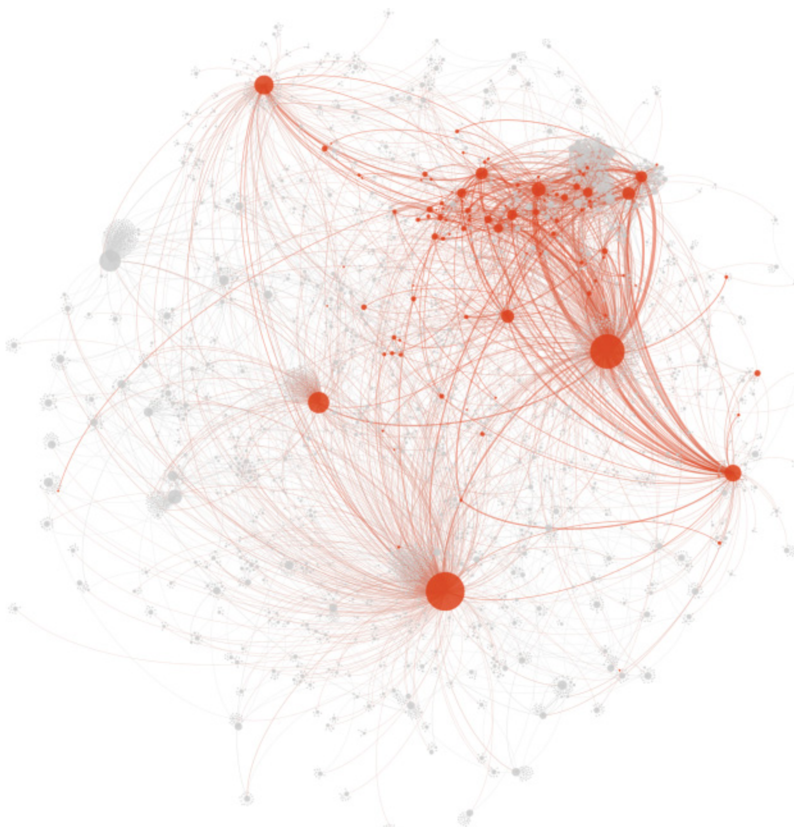

Fig. 14. Mobile ecosystem filtered by overlapping data.

identifiers or naming conventions. Consequently, matching names and labels of firms or individuals across datasets is not a trivial task. Firm names may be inconsistent and use different enterprise labeling. As a result of mergers, acquisitions, or corporate restructuring, firms may also change their names over time. Appropriate identification and matching algorithms must therefore be developed to ensure consistency across datasets.

The use of socially constructed data also has pros and cons. Advantages include its open access and availability, potentially large coverage, timeliness, and community verification of data quality. Some of the disadvantages are the potential of incompleteness and inconsistencies, lack of established perspective, and the issue (although slightly different from that of curated data) of incompleteness and inconsistencies.

Another challenge is the selection of companies and their assignment to market segments (e.g., mobile). Various industry classifications exist, but datasets often use different classification schemes. The identification of primary and secondary market segments is particularly challenging for large firms that operate in multiple and equally important segments. Intelligent market segment identification and assignment methods must therefore be developed. As firms enter, transform, and exit the ecosystem, it is critical to devise appropriate data persistency protocols by identifying events by time and actors involved.

The emergence of high-performance computing tools and the ubiquity of data, however, afford new possibilities. The process of turning data into a dynamic graph or visual model was previously expensive, technically challenging, and regarded more of an experiment's end product than an everyday tool [Fox and Hendler 2011]. The development of visualization technologies has been compounded by the difficulties of introducing new organizational and management processes, which, if not developed, implemented, or applied correctly, may lead to inconclusive results. Particularly in visualizing temporal changes of business ecosystems, node-link configurations are not necessarily unique and results may be misleading. The boundary-setting problem, or inclusion of nodes, is often defined with artificial rules. Sense making and storytelling are thus also a critical step in creating new insights based on the visualizations in data-driven studies.

A much deeper quantitative analysis is possible with multilevel datasets such as the ones reported in this study. We suggest only a few of the many possibilities. As a next step, the data can be further analyzed for predictive analysis, identifying opportunity-based antecedents (such as the likelihood that two nodes will come into contact) and benefit-based antecedents (reflecting utility maximization or discomfort minimization that leads to link formation), along the distinction proposed by Borgatti et al. [2009]. In a similar manner, the descriptive and predictive analysis of link dissolution offers a rich opportunity for further study.

Additionally, a complementary qualitative analysis could study the character of the links and of the players. For example, types and values of deals between organizations could be depicted through link color, line, and direction. Social and cultural characteristics of individuals such as alumni relationships with educational institutions or professional organizations could add information about shared values among cohort groups, perhaps indicating their varying ability to correctly perceive the network around them in order to achieve various outcomes. The visual representation of such qualifications could be integrated into the network visualizations. Additional data sources, such as Twitter, LinkedIn, and patent data, may further enhance the construction of a complete ecosystem perspective. Further, the analysis of network dynamics could provide insights about the antecedents and influences of ecosystem transformation through animated time series showing the creation and dissolution of links enabling the flow of information, talent, and financial resources from one node to another.

There are also many opportunities for creating other representations of business ecosystem dynamics. This may include the development of an interactive visualization system using multiple views. A high granularity in temporal sampling, for example, may mask some key events in ecosystem evolution. The alignment and representation of time units at potentially different scales is an important representational aspect. While established datasets may capture large, less frequent events, socially constructed and curated data may capture activities that occur in closer time intervals. Enabling a user-driven selection of time units will enable greater insight and discovery of the temporal nature of ecosystem activities.

6. CONCLUSIONS

This article advocates a data-driven visualization methodology for analyzing the dynamics of business ecosystems. This approach was illustrated with an exploratory analysis of recent activities of three interfirm subnetworks involving key platform providers Microsoft/Nokia, Google/Motorola Mobility, and Apple/Samsung using three disconnected but highly complementary data sources. Our results show that each dataset contributes differently to synergistic insights and can be used jointly to reveal consistent patterns and create interrelated insights. Visualization of the analysis permits examination of the structure of the ecosystem, in holistic as well as cluster-specific perspectives, each of which can be used to drive their own set of hypotheses. This example of a data-driven analysis provided important insights into patterns of event sequences between nodes for the ecosystem as a whole as well as for three particular subnetworks.

This study introduces a set of lenses to explore many interesting business ecosystem issues including what relationship configurations characterize growth; how the position and role of firms in the ecosystem influence their access to talent, information, and resources; what event windows and types are relevant for observing ecosystem dynamics; and what sequences matter.

Causality in social science research is difficult to establish; this is certainly true of network analysis. Even among documented relationships, it cannot be assumed that a singular event/relationship caused an activity/impact. Specific nodes and links were certainly of interest in this study. However, patterns in the network were the primary focus. Through multiple datasets and a method of triangulating against key nodes, insights were validated and enhanced. In their review of factors leading to emergence in organizations, Padgett and Powell [2012] emphasize the antecedent role of network changes within various levels of a system as well as the critical role of new link formation between nodes in different levels of a system, solving problems by reaching into a different network, as indicators of major reorganizations in networks. Our use of multiple datasets reflecting different levels of the mobile business ecosystem illustrate the benefits of this approach.

The data analysis and visualizations of this study provided a network description of a business ecosystem. Links between organizations in this ecosystem are residual sets of boundaries that reflect relationships that have been created. In their autocatalytic theory of organizational analysis, Padgett and Powell [2012] suggest that in the short run, actors make relationships, while in the long run, relationships make the actors. As illustrated by the analyses of the mobile ecosystem in this study, we suggest that in the transformational flow of transactions, information, talent, and financial resources, personal relationships may be seen as process indicators, deals and events as culmination of those processes.

Conclusions based on these models must thus be carefully scrutinized for the possibility of alternative explanations. Along the same lines, the amount of information that is captured and presented can often be overwhelming to the end-user. In many instances, what ecosystem data is visualized and how depend not only on the nature of the data but also on the question that is being asked, the context in which the question is being asked, and ultimately the cognitive abilities of the user. In order to overcome the aforementioned challenges, researchers must therefore ensure a balance between detail, abstraction, accuracy, efficiency, perceptual tension, and aesthetics in their complex network visualizations [Segel and Heer 2010]. These observations highlight the importance of setting the context and defining the elements very carefully in an ecosystem visualization study.

REFERENCES

R. Adner. 2012. *The Wide Lens: A New Strategy for Innovation*. Penguin Group, New York, NY.

G. Ahuja. 2000. Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly* 45, 3 (2000), 425–455.

G. Ahuja, G. Soda, and A. Zaheer. 2011. The genesis and dynamics of organizational networks. *Organization Science* 23, 2 (2011), 443–448.

B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers. 2014. Visualizing sets and set-typed data: State-of-the-art and future challenges. *EuroVis - STARs*, R. Borgo, R. Maciejewski, and I. Viola (Eds.). The Eurographics Association. DOI: 10.2312/eurovisstar.20141170

B. N. Anand and T. Khanna. 2000. Do firms learn to create value? The case of alliances. *Strategic Management Journal* 21, 3 (2000), 295–315.

A.-L. Barabási. 2003. *Linked: How Everything Is Connected to Everything Else and What It Means*. Plume, New York, NY.

S. J. Barnes and R. T. Vidgen. 2006. Data triangulation and web quality metrics: A case study in e-government. *Information & Management* 43, 6 (2006), 767–777.

R. C. Basole. 2009. Visualization of interfirm relations in a converging mobile ecosystem. *Journal of Information Technology* 24, 2 (2009), 144–159.

R. C. Basole. 2014. Visual business ecosystem intelligence: Lessons from the field. *IEEE Computer Graphics and Applications* 34, 5 (2014), 51–58.

R. C. Basole, T. Clear, M. Hu, H. Mehrotra, and J. Stasko. 2013. Understanding interfirm relationships in business ecosystems with interactive visualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2526–2535.

R. C. Basole, M. Hu, P. Patel, and J. T. Stasko. 2012. Visual analytics for converging-business-ecosystem intelligence. *IEEE Computer Graphics and Applications* 32, 1 (2012), 92–96.

R. C. Basole and W. B. Rouse. 2008. Complexity of service value networks: Conceptualization and empirical investigation. *IBM Systems Journal* 47, 1 (2008), 53–70.

R. C. Basole, C. D. Seuss, and W. B. Rouse. 2013. IT innovation adoption by enterprises: Knowledge discovery through text analytics. *Decision Support Systems* 54, 2 (2013), 1044–1054.

M. Bastian, S. Heymann, and M. Jacomy. 2009. Gephi: An open source software for exploring and manipulating networks. In *Proceedings of the 3rd International AAAI Conference on Weblogs and Social Media*.

L. Bizzi and A. Langley. 2012. Studying processes in and around networks. *Industrial Marketing Management* 41, 2 (2012), 224–234.

N. W. H. Blaikie. 1991. A critique of the use of triangulation in social research. *Quality & Quantity* 25, 2 (1991), 115–136.

S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca. 2009. Network analysis in the social sciences. *Science* 323, 5916 (2009), 892–895.

C. L. Borgman. 2007. *Scholarship in the Digital Age*. MIT Press, Cambridge, MA.

L. Bratthall and M. Jørgensen. 2002. Can you trust a single data source exploratory software engineering case study? *Empirical Software Engineering* 7, 1 (2002), 9–26.

M. Buchanan. 2002. *Small World: Uncovering Nature's Hidden Networks*. Weidenfeld & Nicolson, London, UK.

R. S. Burt. 2009. *Structural Holes: The Social Structure of Competition*. Harvard University Press, Boston, MA.

S. K. Card, J. D. Mackinlay, and B. Shneiderman. 1999. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann.

D. Carson, A. Gilmore, C. Perry, and K. Gronhaug. 2001. *Qualitative Marketing Research*. Sage.

N. K. Denzin. 2009. *The Research Act: A Theoretical Introduction to Sociological Methods*. Aldine De Gruyter.

C. Dhanaraj and A. Parkhe. 2006. Orchestrating innovation networks. *Academy of Management Review* 31, 3 (2006), 659–669.

D. Dougherty and D. D. Dunne. 2011. Organizing ecologies of complex innovation. *Organization Science* 22, 5 (2011), 1214–1223.

K. M. Eisenhardt and C. B. Schoonhoven. 1996. Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms. *Organization Science* 7, 2 (1996), 136–150.

P. Fox and J. Hendler. 2011. Changing the equation on scientific data visualization. *Science* 331, 6018 (2011), 705–708.

M. J. Gallivan. 1997. Value in triangulation: A comparison of two approaches for combining qualitative and quantitative methods. In *Proceedings of the IFIP TC8 WG 8.2 International Conference on Information Systems and Qualitative Research*. 417–443.

M. Gibbert, W. Ruigrok, and B. Wicki. 2008. What passes as a rigorous case study? *Strategic Management Journal* 29, 13 (2008), 1465–1474.

J. Giles. 2005. Internet encyclopaedias go head to head. *Nature* 438, 1 (2005), 900–901.

M. Gladwell. 2000. *The Tipping Point: How Little Things Can Make a Big Difference*. Little, Brown.

D. G. Green and S. Sadedin. 2005. Interactions matter complexity in landscapes and ecosystems. *Ecological Complexity* 2, 2 (2005), 117–130.

R. Gulati. 1998. Alliances and networks. *Strategic Management Journal* 19, 4 (1998), 293–317.

R. Gulati, N. Nohria, and A. Zaheer. 2000. Strategic networks. *Strategic Management Journal* 21, 3 (2000), 203–215.

J. Hagel and J. S. Brown. 2005. *The Only Sustainable Edge: Why Business Strategy Depends on Productive Friction and Dynamic Specialization*. Harvard Business Press.

M. Hammersley and P. Atkinson. 1983. *Ethnography: Principles in Practice*. Routledge, London, UK.

D. Hsu. 2006. Venture capitalists and cooperative start-up commercialization strategy. *Management Science* 52, 2 (2006), 204–219.

M. Huisman and M. A. J. van Duijn. 2005. Software for statistical analysis of social networks. *IEEE Intelligent Systems* 20, 1 (2005), 80–82.

V. W. Hwang and G. Horowitt. 2012. *The Rainforest: The Secret to Building the Next Silicon Valley*. Regenwald.

M. Iansiti and R. Levien. 2004. *The Keystone Advantage: What New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability*. Harvard Business School Press, Boston, MA.

M. Iansiti and G. L. Richards. 2006. The information technology ecosystem: Structure, health and performance. *Antitrust Bulletin* 51, 1 (2006), 77–110.

H. Ibarra and M. T. Hansen. 2011. Are you a collaborative leader? *Harvard Business Review* 89, 7/8 (2011), 68–74.

B. Iyer, C. H. Lee, and N. Venkatraman. 2006. Managing in a small world ecosystem: Some lessons from the software sector. *California Management Review* 48, 3 (2006), 28–47.

T. D. Jick. 1979. Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly* 24, 4 (1979), 602–611.

B. Kaplan and D. Duchon. 1988. Combining qualitative and quantitative methods in information systems research: A case study. *MIS Quarterly* 12, 4 (1988), 571–586.

G. Klein, B. Moon, and R. R. Hoffman. 2006. Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems* 21, 5 (2006), 88–92.

A. Langley. 1999. Strategies for theorizing from process data. *Academy of Management Review* 24, 4 (1999), 691–710.

R. B. Laughlin and D. Pines. 2000. The theory of everything. *Proceedings of the National Academy of Sciences of the United States of America* 97, 1 (2000), 28.

J. Leskovec, D. Huttenlocker, and J. Kleinberg. 2010. Governance in social media: A case study of the Wikipedia promotion process. In *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*.

P. Lord and A. Macdonald. 2003. *e-Science Curation Report: Data Curation for e-Science in the UK: An Audit to Establish Requirements for Future Curation and Provision*. Digital Archiving Consultancy Limited.

K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis. 2012. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications* 32, 1 (2012), 12–19.

M. L. Markus. 1994. Electronic mail as the medium of managerial choice. *Organization Science* 5, 4 (1994), 502–527.

A. McAfee and E. Brynjolfsson. 2012. Big data: The management revolution. *Harvard Business Review* 90, 10 (2012), 60–128.

J. F. Moore. 1993. Predators and prey: A new ecology of competition. *Harvard Business Review* 71, 3 (1993), 75–86.

I. Newman and C. S. Ridenour. 1998. *Qualitative-Quantitative Research Methodology: Exploring the Interactive Continuum*. Southern Illinois University Press.

C. North. 2006. Toward measuring visualization insight. *IEEE Computer Graphics and Applications* 26, 3 (2006), 6–9.

C. Oliver. 1990. Determinants of interorganizational relationships: Integration and future directions. *Academy of Management Review* 15, 2 (1990), 241–265.

A. J. Onwuegbuzie and N. L. Leech. 2007. Validity and qualitative research: An oxymoron? *Quality & Quantity* 41, 2 (2007), 233–249.

J. Owen-Smith and W. W. Powell. 2004. Knowledge networks as channels and conduits: The effects of spillovers in the boston biotechnology community. *Organization Science* 15, 1 (2004), 5–21.

J. F. Padgett and W. W. Powell. 2012. *The Emergence of Organizations and Markets*. Princeton University Press, Princeton, NJ.

M. Q. Patton. 2001. *Qualitative Research & Evaluation Methods*. Sage Publications.

W. W. Powell and E. Giannella. 2009. *Collective Invention and Inventor Networks*. Elsevier, Amsterdam, Netherlands.

I. Prigogine. 1997. *The End of Certainty*. Free Press.

N. Rubens, K. Still, J. Huhtamäki, and M. G. Russell. 2010. Leveraging social media for analysis of innovation players and their moves. Technical report, Media X, Stanford University.

N. Rubens, K. Still, J. Huhtamäki, and M. G. Russell. 2011. A network analysis of investment firms as resource routers in the chinese innovation ecosystem. *Journal of Software* 6, 9 (2011), 1737–1745.

M. G. Russell, K. Still, J. Huhtamäki, C. Yu, and N. Rubens. 2011. Transforming innovation ecosystems through shared vision and network orchestration. In *Proceedings of the Triple Helix IX International Conference*.

R. C. Sampson. 2004. Organizational choice in R&D alliances: Knowledge-based and transaction cost perspectives. *Managerial and Decision Economics* 25, 6–7 (2004), 421–436.

M. A. Schilling. 2009. Understanding the alliance data. *Strategic Management Journal* 30, 3 (2009), 233–260.

M. A. Schilling and C. C. Phelps. 2007. Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science* 53, 7 (2007), 1113–1126.

E. Segel and J. Heer. 2010. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (2010), 1139–1148.

C. D. Seuss. 2009. Meaning extraction helps develop in-depth understanding. *Competitive Intelligence Magazine* 12, 5/6 (2009), 51–54.

C. D. Seuss. 2011. In-depth understanding: Teaching search engines to interpret meaning [Point of View]. *Proceedings of the IEEE* 99, 4 (2011), 531–535.

P. Smagorinsky. 1995. The social construction of data: Methodological problems of investigating learning in the zone of proximal development. *Review of Educational Research* 65, 3 (1995), 191–212.

H. W. Smith. 1975. *Strategies of Social Research: The Methodological Imagination*. Prentice-Hall, Englewood Cliffs, NJ.

K. Still, J. Huhtamäki, M. G. Russell, and N. Rubens. 2012. Paradigm shift in innovation indicators: From analog to digital. In *Proceedings of the 5th ISPIM Innovation Symposium*.

J. J. Thomas and K. A. Cook. 2006. A visual analytics agenda. *IEEE Computer Graphics and Applications* 26, 1 (2006), 10–13.

E. Tufte. 1983. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT.

J. Urry. 2004. Small worlds and the new social physics. *Global Networks* 4, 2 (2004), 109–130.

S. L. Vargo and R. F. Lusch. 2004. Evolving to a new dominant logic for marketing. *Journal of Marketing* 68, 1 (2004), 1–17.

N. Venkatraman and V. Ramanujam. 1987. Measurement of business economic performance: An examination of method convergence. *Journal of Management* 13, 1 (1987), 109–122.

P. C. Wong and J. Thomas. 2004. Visual analytics. *IEEE Computer Graphics and Applications* 24, 5 (2004), 20–21.

S. Xu, X. Chen, and D. Liu. 2009. Classifying software visualization tools using the Bloom's taxonomy of cognitive domain. In *Proceedings of the Canadian Conference on Electrical and Computer Engineering, 2009 (CCECE'09).* IEEE, 13–18.

R. K. Yin. 2008. *Case Study Research: Design and Methods*. Vol. 5. Sage Publications.

A. Zaheer, R. Gzbyk, and H. Milanov. 2010. Its the connections: The network perspective in interorganizational research. *Academy of Management Perspectives* 24, 1 (2010), 62–77.

Received November 2013; revised August 2014; accepted January 2015