Big Data Analytics: Perspective Shifting from Transactions to Ecosystems

Daniel Zeng, University of Arizona and Chinese Academy of Sciences
Robert Lusch, University of Arizona

Editors in chief (EICs) need to perform a variety of tasks. They shepherd the review process for each and every submitted manuscript, ensuring fairness and timeliness of the review, and technical quality of accepted work. They work closely with the editorial boards and staff to come up with interesting and meaningful topics for special issues. They worry about the academic standing of the publications they oversee, and along with it, a handful of seemingly scientific and quantifiable measures such as impact factors, with changes about how to calculate them imposed by commercial entities from time to time.

They also must consider changes in academic publishing such as open access, social media, and to a lesser degree the business side of publication operations. But above all, as we’re sure that most current and past EICs would agree, one of the most daunting aspects of serving as the EIC is to produce editorials or editor’s letters, just like the one you’re reading right now, on a regular basis. For IEEE Intelligent Systems, this happens fairly frequently—every two months. According to a recent readership survey, IS readers actually have high expectations of such editorials, preferring informative content with broad implications. This finding doesn’t exactly ease the stress associated with editorial writing. As a researcher, any EIC is bound by the scope of his or her own areas. However, IS readers made it known that they want to read about topics beyond the EIC’s own areas. It seems clear that the old model of having the EIC be solely responsible for editorials needs updating.

So, with the endorsement of the editorial board, starting in this issue, we’ll experiment with a new hybrid model of editorial writing. Every year, we’ll reserve several “A Letter from the Editor” slots to be co-authored by the EIC and one or two guest researchers. These co-authored pieces will focus on certain areas of great promise in AI, raising their awareness and summarizing major challenges. In these pieces, we’ll also communicate with readers about what IS is doing in these areas. Then we’ll formally evaluate the perceived value of this new format in early 2014 through a readership survey.

The current editor’s letter is co-authored by Robert Lusch, a former EIC for the Journal of Marketing, one of the top marketing publications. Lusch is well known for co-developing one of the theoretical milestones in the field of service sciences and systems—service-dominant logic (S-D Logic; www.sdlogic.net). He’s no stranger to IS; he served as...
a guest editor for our November/December 2010 special issue on Social Media Analytics and Intelligence. Here, we’ll discuss how to better understand and make use of big data analytics in today’s business environment through a new ecosystem-based perspective.

**Big Data Analytics**

Across disciplines, big data has been attracting significant attention globally from government funding agencies, academia, and industry. The field of AI is no exception, with its particular emphasis on developing specialized data-mining methods to explore big data, among other closely related research topics that can be broadly labeled as *analytics*.

Recognizing the practical relevance of and significant technical challenges associated with big data analytics, *IS* is actively pursuing two projects. First, through the newly established “AI Innovation in Industry” department, under the leadership of Daniel O’Leary, we’re working on a series of “big data analytics practice and innovation” department articles, in collaboration with a select set of industry research centers and IT departments. These industry units cover both the supply and demand side of the emerging big data analytics industry. This department’s objective is to sample best practices in big data analytics; catalog key technical challenges; identify the gaps between theory, technology, and practice; and map out near-future directions from an industry perspective. Second, preparation for a Big Data Analytics special issue is under way. We expect to publish this special issue in Fall 2014.

**Towards an Ecosystem-Based Perspective**

It’s interesting to consider the current knowledge-intensive service economy, and how that affects the way that we should perceive, analyze, and use big data. Even before big data entered our societal consciousness and became a household expression, a dedicated IT sector focusing on delivering data and analytics services had come into being and has been playing an increasingly important role in enabling change in business and government operations. Big data highlights the fact that the scale and variety of potentially relevant data—coupled with real-time, rapid generation of such data—have reached a point at which known computational methods can’t effectively function. Technical literature abounds with ideas on how to deal with such challenges.

Here, however, we argue that beyond technical advances, which are largely about answering the same old questions at the scale and complexity of big data, significant research and application opportunities exist to ask and answer new questions. Realizing such opportunities amounts to (potentially) drastic departures from the existing analytics framework, which is largely *transaction-based*. Consider business applications of data mining and analytics in general—such techniques have been applied with great success to various business operational scenarios (such as customer segmentation and targeting, management of manufacturing systems, and financial and accounting activities) in which organizational structures are fixed and known, and operations are typically designed and implemented in a top-down fashion with well-defined actors, clear boundaries, explicit communication channels, and a rigid structure. Developing a new computing infrastructure (cloud-oriented hardware/operating systems/file systems, customized database management systems, and specialized query languages) and new algorithms to handle transaction-based analyses and exploration for big data-era applications is certainly important, and has been the norm of recent big data...
analytics research. A case in point would be perabyte-scale weblog analysis, an application about which a wide arrange of e-commerce and social networking companies deeply care.

However, this is only half of the story. Partially driven by wide adaption of Web technology, mobile devices and services, and embedded systems, the past decade has witnessed the birth of a new breed of adaptive enterprises and value networks that rely on data and analytics. The marketplace is changing rapidly. New markets are being created constantly and old ones die. New alliances among upstream and downstream firms and even competitors are forged rapidly and often dissolve quickly. More industries are using information sharing and interlocked collaborative decision making across organizational boundaries. In short, the boundaries between designing, producing, and consuming are blurred. The constant cycle of sense-respond-adapt and bottom-up interconnectivity is prevalent.

At the macroscopic level, the entire industrial landscape is moving away from rigid structures. Instead, we’re seeing less structured, but co-dependent and co-evolving ecosystems. In this new norm of business operation, transactions often take place in a much less-structured environment than before. In a highly structured environment, analyzing transactions in isolation makes sense. In a dynamically evolving ecosystem, however, isolated analyses provide only limited use. Far more important are more holistic, system-level, and integrative analyses. What was provided before as input—the structures—could well be output in the emerging ecosystem-based analysis framework. Another way of saying this is that we must develop analytical tools along with big data to see how structures emerge, proliferate, and morph into other structures.

Think of it this way: a market isn’t a static, rigid structure that can be segmented; rather, a market emerges, and as it grows it splinters into segments. However, just as the market isn’t static, the segments are dynamic and can be viewed as communities of actors that themselves co-evolve—and in the process, they influence other actors and markets.

This ecosystem perspective is motivated and necessitated by practice. Technically, assuming this perspective—or at least taking this as a complementary perspective of the more traditional transaction-based perspective—can lead to fresh insights. But just what is this ecosystem to which we refer?

Think of it as largely and loosely coupled actors that sense, respond, and adapt to each other through a common language, technology, and institutions to engage in an exchange of some sort. This is often economic, but also social exchange. Similar to biological ecosystems, the actors involved can become extinct or thrive and flourish. In virtually all cases they’re trying to solve local problems and adapt to local conditions. As with biological ecosystems, the actors exchange—and this is increasingly the exchange of some service that’s enabled by applying knowledge.

As we stated at the outset, this is the basis of a knowledge-intensive service economy that’s engulfing nations and the world. The actors are humans or groups of them, such as enterprises, but also the actors can be computerized systems. These ecosystems can be thought of in terms of industries, but that’s increasingly myopic, because with digitization and other technologies, the lines where industries start and end are less clear. More importantly, industries themselves are defined around units of output, sales transactions, and market shares, which is the problem we’re trying to overcome by thinking in terms of ecosystems. We don’t think of industries in terms of units of output, but rather the processes and flows that reflect the service they provide. For instance, the elevator business becomes the people and material flow business or the pharmaceutical business becomes the enhancement of human capabilities business.

Once we move beyond data focused on transactions and direct exchanges, we begin to see how the ecosystem assembles and disassembles itself. And as we observe this process, we’re able to begin building analytics that model assembly rules, which helps us build artificial worlds to mirror the actual world. Experimentation then moves out of the real world and into the artificial world.

This brings us back to a part of ecosystems that we haven’t discussed: institutions. Institutions are clever human inventions that let us coordinate their thinking and behavior. Some of these are explicit, such as written laws, but many are unwritten and tacit, such as how actors are polite or impolite to other actors based on social norms or how they’re honest or attempt to deceive. Of course, it isn’t that simple, because not all actors are governed or controlled by the same institutions. However, big data with an ecosystem lens lets us use computational means to identify overlapping institutional structures or discover how some institutions are nested within other larger (macro) institutions. The challenge strategists and policy makers have in terms of institutions is that they change quite slowly, and it’s hard to experiment with different institutions, except over long time frames. However, if we take an ecosystems view of big data and build artificial worlds in which we can manipulate institutions, we can provide an entire host of analytics to the policy maker.

A ecosystem-based perspective has the potential of helping define
new classes of data mining and analytics problems. Compared to transaction analyses, big data perhaps plays a more active role in this ecosystem-based framework, as more global perspectives and less structure entails data from a larger context. An ecosystem-based framework needs analyses done at multiple levels of abstraction, something right in the AI domain. It also needs analysis that can truly handle multiagent contexts and complications, for which AI has a lot to offer. We’re eager to see the emergence of this new breed of ecosystem-based big data analytics work, rooted in real practice with major AI contributions.

Acknowledgments
This work is partially supported by the National Natural Science Foundation of China Grants 71025001 and 70890084. Daniel Zeng is also affiliated with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences.

New Editorial Board Members

Bhavani Thuraisingham is the Louis A. Beecherl, Jr. Distinguished Professor of Computer Science and the executive director of the Cyber Security Research and Education Institute (CSRI) at The University of Texas at Dallas. Her research interests include data security, data mining, and social computing. Thuraisingham has a PhD in theory of computation from the University of Wales, UK. She is an IEEE Fellow and received the 1997 IEEE Computer Society Technical Achievement Award.

Longbing Cao is a professor of information technology at the University of Technology, Sydney, Australia, where he is also the founding director of the university’s Advanced Analytics Institute. His main research interests include data mining and knowledge discovery, machine learning, behavior informatics, agent mining, multiagent systems, and open complex intelligent systems. Cao has PhDs in intelligent sciences and computing sciences, and has been leading many enterprise projects in advanced analytics, data mining, and behavior study.

Daniel B. Neill is the Dean’s Career Development Professor and associate professor of information systems at Carnegie Mellon University’s Heinz College, where he directs the Event and Pattern Detection Laboratory. His research interests include machine learning, data mining, and event detection in massive datasets. Neill has a PhD in computer science from Carnegie Mellon University.

Lefei Li is an associate professor in the Department of Industrial Engineering at Tsinghua University. His research interests include intelligent transportation systems (ITS), logistics and service operations management, and complex systems modeling and simulation. Li has a PhD in systems and industrial engineering from the University of Arizona.

IEEE Computer Society is offering $40,000 in student scholarships, from $1,000 and up, to recognize and reward active student volunteer leaders who show promise in their academic and professional efforts.

Graduate students and undergraduate students in their final two years, enrolled in a program in electrical or computer engineering, computer science, information technology, or a well-defined computer-related field, are eligible. IEEE Computer Society student membership is required.

Apply now! Application deadline is 30 September 2013. For more information, go to www.computer.org/scholarships, or email patricia.edwards@computer.org.

To join IEEE Computer Society, visit www.computer.org/membership.