COMPUTING edge

END-OF-YEAR WRAP-UP

Also in this issue:

> Extending Our Field's Reach

> Multimedia Big Data Computing

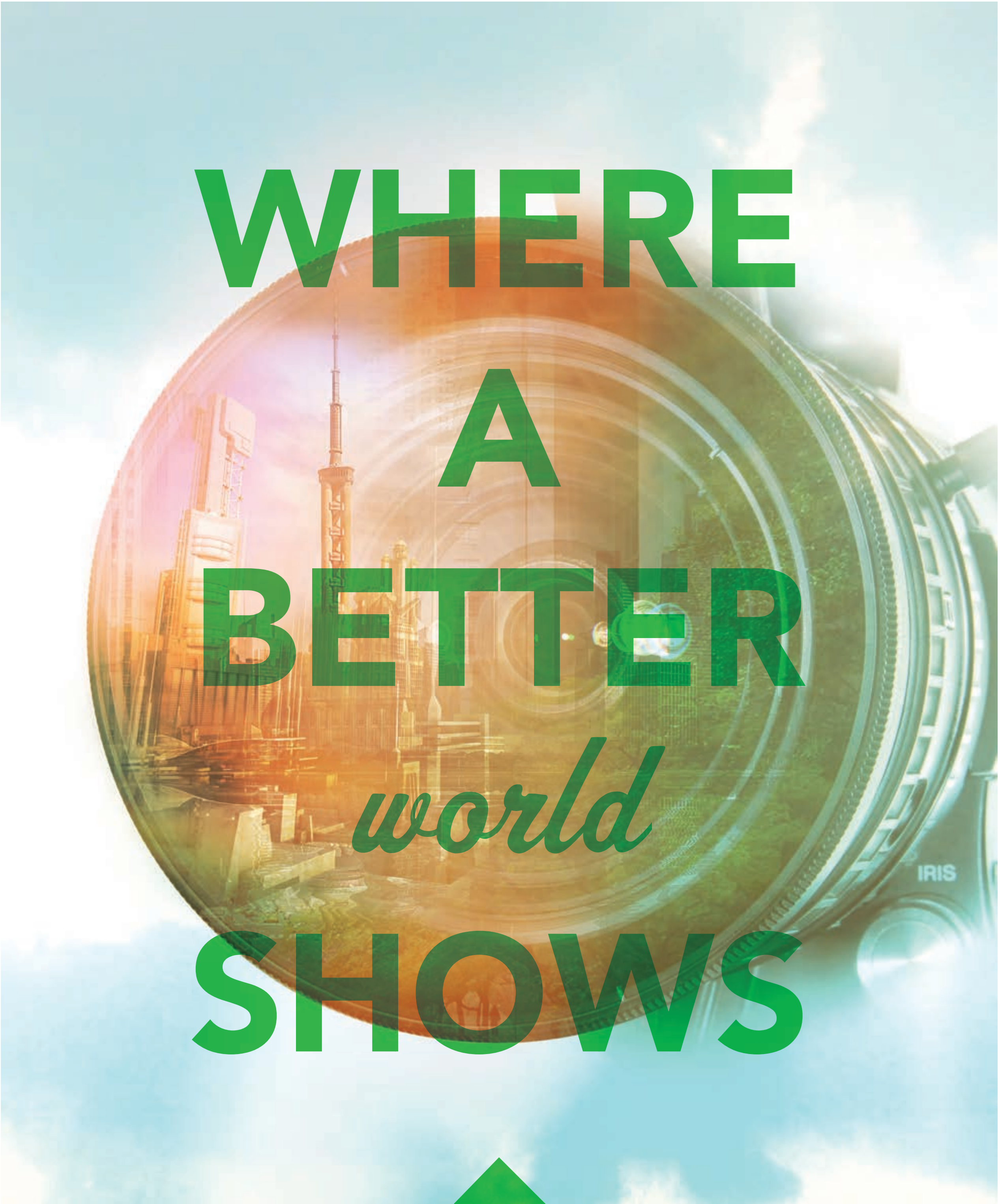
DECEMBER 2015

www.computer.org

IEEE

IEEE computer society

#CES2016



CES IS FAR MORE THAN COOL TECH. IT'S THE TECHNOLOGY AND IDEAS THAT ARE BRINGING NEW OPPORTUNITY AND A BETTER LIFE TO DEVELOPING COUNTRIES AND BEYOND. REGISTER TODAY.

CES® 2016 JAN. 6-9, 2016

TECH EAST • TECH WEST • TECH SOUTH
LAS VEGAS, NV

REGISTER NOW at **CESweb.org**

CES

THE GLOBAL STAGE FOR INNOVATION

PRODUCED BY CEA

COMPUTING edge

IEEE COMPUTER SOCIETY http://computer.org • +1 714 821 8380

Printed with inks containing soy and/or vegetable oils

IEEE

IEEE computer society

STAFF

Editor
Lee Garber

Contributing Editors
Christine Anthony, Brian Brannon, Carrie Clark Walsh, Chris Nelson, Meghan O'Dell, Dennis Taylor, Bonnie Wylie

Production & Design
Carmen Flores-Garvey, Monette Velasco, Jennie Zhu-Mai, Mark Bartosik

Manager, Editorial Services Content Development
Richard Park

Senior Manager, Editorial Services
Robin Baldwin

Director, Products and Services
Evan Butterfield

Senior Advertising Coordinator
Debbie Sims

BPA WORLDWIDE

Circulation: ComputingEdge is published monthly by the IEEE Computer Society. IEEE Headquarters, Three Park Avenue, 17th Floor, New York, NY 10016-5997; IEEE Computer Society Publications Office, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720; voice +1 714 821 8380; fax +1 714 821 4010; IEEE Computer Society Headquarters, 2001 L Street NW, Suite 700, Washington, DC 20036.

Postmaster: Send undelivered copies and address changes to ComputingEdge-IEEE Membership Processing Dept., 445 Hoes Lane, Piscataway, NJ 08855. Application to Mail at Periodicals Postage Prices is pending at New York, New York, and at additional mailing offices. Canadian GST #125634188. Canada Post Corporation (Canadian distribution) publications mail agreement number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8 Canada. Printed in USA.

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in ComputingEdge does not necessarily constitute endorsement by the IEEE or the Computer Society. All submissions are subject to editing for style, clarity, and space.

Reuse Rights and Reprint Permissions: Educational or personal use of this material is permitted without fee , provided such use: 1) is not made for profit; 2) includes this notice and a full citation to the original work on the first page of the copy; and 3) does not imply IEEE endorsement of any third-party products or services. Authors and their companies are permitted to post the accepted version of IEEE-copyrighted material on their own Web servers without permission, provided that the IEEE copyright notice and a full citation to the original work appear on the first scree n of the posted copy. An accepted manuscript is a version which has been revised by the author to incorporate review suggestions, but not the published version with copy-editing, proofreading, and formatting added by IEEE. For more information, please go to: http://www.ieee .org/publications_standards/publications/rights/paperversionpolicy.html. Permission to reprint/republish this material for commercial, advertising, or promotional purposes or for creating new collective works for resale or redistribution must be obtained from IEEE by writing to the IEEE Intellectual Property Rights Office, 445 Hoes Lane, Piscataway, NJ 08854-4141 or pubs-permissions@ieee .org. Copyright © 2015 IEEE. All rights reserved.

Abstracting and Library Use: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Unsubscribe: If you no longer wish to receive this *ComputingEdge* mailing, please email IEEE Computer Society Customer Service at help@computer.org and type "unsubscribe ComputingEdge" in your subject line.

IEEE prohibits discrimination, harassment, and bullying. For more information, visit www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

IEEE Computer Society Magazine Editors in Chief

Computer
Sumi Helal, *University of Florida*

IEEE Software
Diomidis Spinellis, *Athens University of Economics and Business*

IEEE Internet Computing
M. Brian Blake, *University of Miami*

IT Professional
San Murugesan, *BRITE Professional Services*

IEEE Security & Privacy
Ahmad-Reza Sadeghi, *Technical University of Darmstadt*

IEEE Micro
Lieven Eeckhout, *Ghent University*

IEEE Computer Graphics and Applications
L. Miguel Encarnação, *ACT, Inc.*

IEEE Pervasive Computing
Maria Ebling, *IBM T.J. Watson Research Center*

Computing in Science & Engineering
George K. Thiruvathukal, *Loyola University Chicago*

IEEE Intelligent Systems
Daniel Zeng, *University of Arizona*

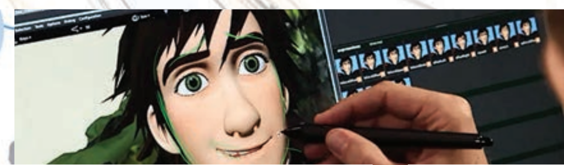
IEEE MultiMedia
Yong Rui, *Microsoft Research*

IEEE Annals of the History of Computing
Nathan Ensmenger, *Indiana University Bloomington*

IEEE Cloud Computing
Mazin Yousif, *T-Systems International*

DECEMBER 2015 • VOLUME 1, NUMBER 12

COMPUTING edge



42

Multimedia
Big Data
Computing

Departments

SPOTLIGHT ON TRANSACTIONS

Seamless Learning: Using Location-Aware Technology to Support Art Education

Mike Sharples, The Open University

This installment highlighting the work published in IEEE Computer Society journals comes from IEEE Transactions on Learning Technologies.

Schools and universities teach fundamental principles, universal facts, and generalized skills. But some fields also require students to gain knowledge that's bound to locations or artifacts. Becoming an expert in, for example, geology, medicine, civil engineering, art, or history requires linking situated and generalized knowledge. In this way, general principles are applied in various specific settings, while localized practices, such as recognizing specific disease symptoms or appreciating art, form abstract knowledge.

Seamless learning is a new concept proposing that previously distinct learning experiences (inside and outside classrooms; on and off campus) should be linked to appear continuous.[1] Personal location-aware technologies, such as mobile devices, can assist this flow by allowing students to collect learning experiences at home, outdoors, and in enriched locations such as museums or field trips, and then examine these experiences in the classroom.[2]

In a 2014 *IEEE Transactions on Learning Technology* article, Yael Kali and her colleagues at the University of Haifa described their two-year study in which location-aware technology supported art history students' learning in three locations.[3] In the classroom, instructors demonstrated relevant art-appreciation skills using paintings from the course website. The students then visited an art gallery equipped with a Tumblr mobile phone app as well as a custom app that detected the students' locations and offered a short multimedia presentation about the artwork. While in the gallery, student teams, which could still access the course website, used the Tumblr app to document their ideas. They continued this collaborative activity at home through a shared Google document. Lastly, the students submitted their group assignments to the instructor for feedback. Kali and colleagues reviewed student assignments, observed lessons within classrooms, interviewed instructors and students, and administered a student questionnaire.

In using a design-research approach to developing and evaluating technology-enhanced learning, Kali and her team had dual objectives: to improve the learning system's design and to observe how the learning was enacted in practice. As the activity progressed from the classroom to the museum to the students' homes, the instructor's contributions decreased, supporting both independent and collaborative learning. Data analyses showed that the course website was key to connecting the students' learning across locations. However, the students perceived the app and museum guide as having low usability. Future work should focus on developing location-aware mobile applications that enhance learning and enable smooth activity transitions across the various settings.

REFERENCES

1. G.D. Kuh, "Guiding Principles for Creating Seamless Learning Environments for Undergraduates," *J. College Student Development*, vol. 37, no. 2, 1996, pp. 135–148.
2. L.-H. Wong and C.K. Looi, "What Seams Do We Remove in Mobile-Assisted Seamless Learning? A Critical Review of the Literature," *J. Computers and Education*, vol. 57, no. 4, 2011, pp. 2364–2381.
3. Y. Kali, O. Sagy, O. Mogilevsky, and E. Maayan-Fanar, "Harnessing Technology for Promoting Undergraduate Art Education: A Novel Model that Streamlines Learning between Classroom, Museum, and Home," *IEEE Trans. Learning Technologies*, vol. 8, no. 1, 2014, pp. 5–16.

MIKE SHARPLES is a professor of educational technology at The Open University. Contact him at mike.sharples@open.ac.uk.

2376-113X/15/$31.00 © 2015 IEEE

CS FOCUS

Magazine Roundup

The IEEE Computer Society's lineup of 13 peer-reviewed technical magazines covers cutting-edge topics ranging from software design and computer graphics to Internet computing and security, from scientific applications and machine intelligence to cloud migration and microchip manufacturing. Here are highlights from recent issues.

Computer

To be explored and understood, major scientific frontiers require continued innovation in high-performance computing. This is the subject of *Computer*'s November 2015 special issue on **grand challenges in scientific computing**.

IEEE Software

Refactoring changes a program's source code without changing its external behavior, typically to improve the design. Articles in *IEEE Software*'s November/December 2015 special issue on refactoring range from the historical, exploring the approach's research origins; to the practical, exploring software developers' experiences with it; to the theoretical, exploring new techniques that haven't been used in production yet.

IEEE Internet Computing

Our ancestors left behind few records, but today, we're creating and preserving increasingly complete digital traces and models of almost every aspect of our lives. *IEEE Internet Computing*'s November/December 2015 special issue on **the Internet of you** explores this phenomenon—from small user-centric models of individuals to real-time analytics of large aggregations of user data—including its technologies and challenges.

Computing in Science & Engineering

Global climate change and its impact on natural resources, infrastructure, and health are among the 21st century's biggest challenges. The climate-science and climate-change fields rely heavily on computationally intensive simulations and data analysis. *CiSE*'s November/December 2015 special issue on **computing and climate** examines these topics.

IEEE Security & Privacy

In response to ongoing challenges, information-security spending has grown steadily and might eventually become more than many organizations can afford. To deal with rising costs, governments and companies must carefully balance tradeoffs between security and privacy. These and related matters are discussed in *IEEE S&P*'s September/October 2015 special issue on the **economics of cybersecurity**.

IEEE Cloud Computing

"**The Strategic Value of the Cloud**," from *IEEE Cloud Computing*'s July/August 2015 issue, identifies four major strategies that exploit the cloud and related approaches such as big data, social media, mobile technology, and the Internet of Things. The article also describes the strategies' high-level architectural patterns.

IEEE Computer Graphics and Applications

The authors of "More Than Telling a Story: Transforming Data into Visually Shared Stories," from *IEEE CG&A*'s September/October 2015 issue, look at how the visualization community has discussed **visual storytelling**. They present a visual-data storytelling process and discuss future research opportunities.

IEEE Intelligent Systems

In this era of big data, **knowledge engineering** faces fundamental challenges caused by fragmented knowledge from heterogeneous, autonomous sources with complex, evolving relationships Older knowledge-representation, -acquisition, and -inference techniques must be updated. In *IEEE Intelligent Systems*' September/October 2015 issue, "Knowledge Engineering with Big Data" presents BigKE, a knowledge-engineering framework that promises to help alleviate this problem.

IEEE MultiMedia

Recent developments in the way photographs are captured offer the possibility of creating new types of visual connections and conversations. "Let's Weave the Visual Web," from IEEE MultiMedia's July–September 2015 issue, reviews visual documentation's evolution, considers where we're headed, and introduces the visual Web.

IEEE Annals of the History of Computing

In the **early days of digital computing**, researchers often used the computation of a list of prime numbers as a test or demonstration problem. The evolution of this practice is addressed in "Computing Primes (1929–1949): Transformations in the Early Days of Digital Computing," from *IEEE Annals*' July–September 2015 issue.

IEEE Pervasive Computing

Technology use is already pervasive across the food industry. For example, sensors help monitor plant and animal health, and many applications focus on the food supply chain, consumer grocery shopping, or individual dietary activity. Technology improvements will only increase this trend, which is explored in *IEEE Pervasive Computing*'s October–December 2015 special issue on **pervasive food**.

IT Professional

The authors of "**Wearable Computing for the Internet of Things**," from *IT Pro*'s September/October 2015 issue, analyze the characteristics of wearable applications for use with the Internet of Things. They also describe the interaction patterns that should occur between wearable or mobile devices and smart objects.

IEEE Micro

As the benefits of Moore's law progressively diminish and come at a greater cost, there is a growing push to consider alternative approaches to increasing computational capabilities fast enough to continue spurring innovation. *IEEE Micro*'s September/October 2015 special issue on **alternative computing designs and technologies** presents several of these approaches.

Computing Now

The Computing Now website (http://computingnow.computer.org) features **up-to-the-minute computing news** and blogs, along with articles ranging from peer-reviewed research to opinion pieces by industry leaders.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.

EDITOR'S NOTE

End-of-Year Technology Wrap-Up

As is always the case in the world of technology, this year has been a busy one. New technologies have become popular, old ones have fallen by the wayside, and approaches that were picking up momentum have gained traction in the marketplace.

This *ComputingEdge* issue highlights some of 2015's most important trends. For example, *IEEE Computer Graphics and Application*'s "Premo: DreamWorks Animation's New Approach to Animation" looks at the famous DreamWorks studio's new state-of-the-art animation platform, which offers a more efficient, engaging, and natural interface than those of other professional tools.

IEEE Cloud Computing's "The Strategic Value of the Cloud" addresses four major strategies that exploit the cloud and related approaches such as big data, social networking, mobile technology, and the Internet of Things. The article also explains each strategy's high-level architectural patterns.

The authors of "Saving Rhinos with Predictive Analytics," from *IEEE Intelligent Systems*, describe their anti-poaching engine, which builds on behavioral models of both rhinos and poachers to protect as many animals as possible.

"The Ecology of Software Ecosystems," from *Computer*, compares software ecosystems with biological ecosystems to yield new strategies for improving the former's effectiveness and resilience.

IEEE Internet Computing's "Smart Cities' Data: Challenges and Opportunities for Semantic Technologies" explores how we can innovate smart systems for smart cities to make data available homogeneously, inexpensively, and flexibly while supporting new applications.

"Children of the Magenta," from *IEEE Security & Privacy*, argues that today's increased cybersecurity automation reduces humans' awareness of and ability to cope with threats, which can cause serious problems.

ComputingEdge articles on other subjects include the following:

- *IEEE Software*'s "Extending Our Field's Reach" contends that the techniques and processes software engineers have perfected to manage complex projects—such as version control and international collaboration—could benefit other fields and industries.
- In *IT Professional*'s "Executive Roundtable Series: Machine Learning and Cognitive Computing," participants in a recent roundtable discuss the two increasingly high-profile fields.
- *IEEE MultiMedia*'s "Multimedia Big Data Computing" talks about the principal challenges of and approaches to processing the large amounts of multimedia data being generated today.

The Strategic Value of the Cloud

A RECENT STUDY BY OXFORD ECONOMICS BASED ON A RESEARCH COLLABORATION WITH SAP IDENTIFIED THE TOP FIVE BENEFITS THAT EXECUTIVES EXPECT FROM CLOUD INVESTMENTS.[1] First, they expect significant gains in productivity—to do more with less. They also expect significant gains in innovation as well as in the speed and efficiency of processes. These executives believe that cloud investments should enable IT to be a better partner to the business, and IT to become a profit center. Only after these five top priorities did executives rank cost savings.

To be sure, the cloud has a variety of mechanisms that can help reduce costs. For example, a reduction in required physical servers achieved through virtualization or containers reduces total capital expenditures, leasing, or reserved or on-demand costs. Pay per use is a pricing mechanism that can reduce total costs when used with pure public or hybrid clouds, especially in the presence of variable demand.[2] The cloud can also improve performance, for example, by reducing latency for interactive tasks through geographic dispersion, and reducing total response time through parallelism.[2]

However, the Oxford Economics study illustrates that executives don't view cost reduction per se as a major driver for cloud. This is a wise assessment. After all, with IT budgets averaging a few percent of revenues, even a dramatic 25 percent cut in budgets would have a small impact on the finances of the overall corporation, especially compared to, say, increasing revenues by 10 percent, or even preserving existing revenues in the face of turbulence and competition. Such a strategic impact is becoming increasingly common: consider the battles between Borders Books and Amazon.com (the retail division of Amazon, not Amazon Web Services, the cloud provider), Netflix and Blockbuster, Uber and taxis, or WhatsApp and SMS services.

This, however, begs the question of exactly how the cloud can be strategic, and whether such strategies are repeatable. I've identified four major generic strategies, which I call *digital disciplines*, that exploit the cloud and related technologies such as big data, social, mobile, and the Internet of Things.[3] Here, I'll also describe the high-level architectural patterns that each strategy entails.

The digital disciplines update a two-decade-old framework, which articulates three strategies—called *value disciplines*—for achieving marketplace success by offering customers unparalleled value: *operational excellence* (that is, better processes), *product leadership* (that is, better products and services), and *customer intimacy* (that is, better customer relationships).[4]

Given today's technologies, however, physical operational excellence must be complemented by virtual *information excellence*. Standalone product (or service) leadership needs to evolve to smart, digital products and services connected to the cloud and from there onward to an ecosystem of services and partners. I call this *solution leadership*. Traditional face-to-face customer intimacy is now not only complemented by virtual customer intimacy through social networks and online communities, but is also being supplanted by *collective intimacy*, where de-

JOE WEINMAN
joeweinman@gmail.com

 2376-113X/15/$31.00 © 2015 IEEE

tailed data from all individuals, such as movie-watching habits or genetic data, is processed to provide personalized, contextualized recommendations and services to every individual. Finally, a fourth metastrategy, *accelerated innovation*, enables companies to innovate not only faster, but with higher quality and at lower cost, through cloud-based approaches such as contests, idea markets, and innovation networks.

Information Excellence

Traditional operational excellence entailed static design of operational processes such as manufacturing, distribution, or service operations. For example, consider a manufacturer who might spend years designing a new automobile assembly plant, paper mill, or refinery. A model of the new paradigm is extreme flexibility coupled with dynamic optimization. A good example is a modern container port, with numerous trucks, ships, cranes, and containers. The objectives are to maximize throughput, minimize delay, and minimize cost in the face of congestion, ship delays, labor stoppages, and congested roads. Or consider changes in distribution. Rather than a fixed delivery route, a package delivery company wants to optimize deliveries by minimizing fuel and carbon footprint and labor costs and capital requirements while maximizing customer satisfaction through a combination of regularity (for example, deliveries to Acme, Inc. are usually at 10 a.m., after the daily status call) and meeting delivery deadlines and customer constraints.

Today's dynamic optimization problems are computationally complex, that is, intractable. But that's only part of the story. Abstractly they're NP-complete, but that assumes that the problem data can be collected and then a solution attempted. In real life, additional problems include acquiring valid data in real time, updating it as conditions change (a delayed ship, a congested traffic route, a truck with a flat tire) then solving it, at least through heuristics and approximation methods, and then implementing the solution (for example, routing trucks, ships, cranes, and so on). Moreover, global policies might need to be implemented. For example, delivery companies such as UPS have long avoided left turns in their route construction to improve productivity, but recently, New York City has requested that Google help reduce left turns for Google Maps users to enhance pedestrian safety.[5] Solving these types of problems requires collecting big data in real time from things and people, processing it in near real time through an optimal combination of edge and cloud, and then enacting the solution through people and things, as shown in Figure 1.

FIGURE 1. Information excellence high-level architecture. Data is aggregated in the cloud, where it's processed, and the solution implemented through people and things.

Solution Leadership

Products like the Rolex Daytona or Aston-Martin DB9 have traditionally been emblematic of product leadership, as have services such as those from Nordstrom or the Four Seasons. However, today, standalone products and services are evolving into smart, digital ones, with embedded CPUs, sensors, and power. Moreover, they're being connected over mobile and wireline networks to the cloud, which can aggregate and process data from multiple endpoints. Sometimes data aggregation is across multiple similar endpoints, as when smart meters collect usage data for demand response in smart grids, or when wind turbines tune themselves based on the performance of nearby turbines. Sometimes data can be aggregated across heterogeneous endpoints from multiple vendors in an authorized or emergent partner product ecosystem, as when

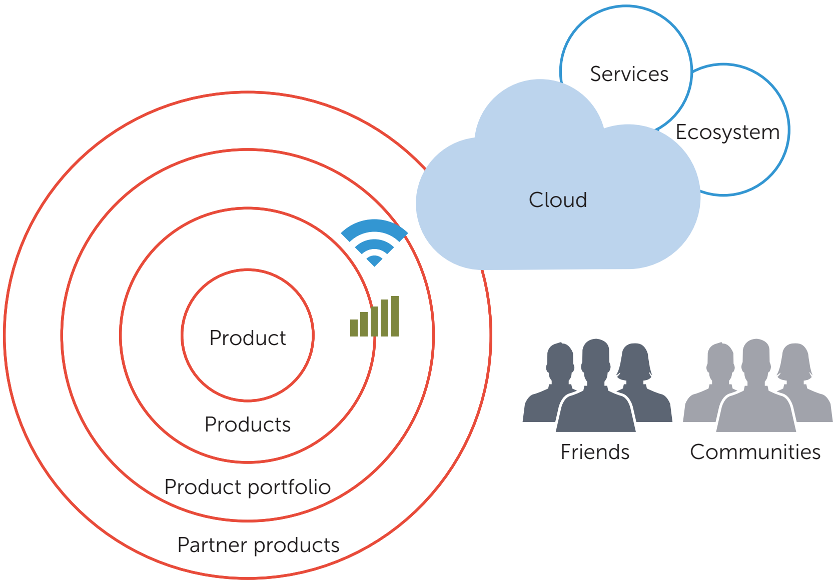


FIGURE 2. Solution leadership high-level architecture. The cloud ties together products, services, consumers, providers, and an extensible partner ecosystem.

activity trackers, footwear, smartphones, and connected scales can interoperate to aid in achieving fitness goals.

Not only products get connected as the Internet of Things connects to the cloud, services do as well. Healthcare services can leverage data from DNA sequencers and CT scanners and connected pacemakers to deliver higher quality services. Fast food restaurants can enable order configuration over the Web, with food deliveries tracked and reported to customers in real time via connected vehicles.

Smart, digital, connected product and service solutions enable ongoing customer relationships, encouraging stickiness and transforming one-time transactions focused on sales to ongoing subscription relationships focused on customer outcomes. The cloud becomes the nexus of data aggregation in real time, the development of and control point for actionable intelligence, and the gateway to social networks, communities, and extensible ecosystems, as Figure 2 illustrates.

Collective Intimacy

The traditional model of customer intimacy is represented by relationships between consumers and their hairstylists, bartenders, butchers, tailors, or physicians, or businesses and the account teams that service them. Such pairwise intimate relationships have moved online to contexts such as social media.

However, there are deeper forces at work, enabled by big data processed by near-infinite cloud resources. Rather than dozens or millions of pairwise intimate relationships, comprehensive data from all customers is collectively analyzed to provide better services to each customer. For example, Netflix collects data on customer *characteristics* (such as demographics and personas) streaming video viewing *behaviors* (rewind, pause, fast forward, watch again, never watch), and *contexts* (device type, time of day, geolocation) from each of their tens of millions of customers.[6] This is combined with external metadata (name, release date, director, locations, and so on) and human-generated tags that assess such elements as the emotional content of the movie (happy, sad, romantic, and so on). All this data is processed to generate personalized recommendations to each viewer. Similarly, the Mayo Clinic processes collective genomic, epigenetic, microbiomic, and pharmacological efficacy data to generate personalized recommendations: personalized medicine and patient-specific therapies.

Here the role of the cloud is to collect, aggregate, and process data, determine personalized services and recommendations, and deliver them to customers, as Figure 3 shows. In addition to behaviors, contexts, tags, and external metadata, social elements can be incorporated as well. For example, Netflix can alter recommendations based on friends' viewing preferences.

Accelerated Innovation

The cloud can also be the means for firms with problems to connect with solvers, through idea markets, challenges, and innovation networks. Traditionally, companies would solve technical challenges internally through R&D labs. The "open innovation" approach espoused by Henry Chesbrough advocates looking beyond the firm boundaries to partners with either technologies or distribution capabilities.7 However, such open innovation is no longer based on intellectual property licensing and static agreements, but can be conducted in a highly dynamic, ad hoc fashion.

One way to do this is through challenges and contests such as the Netflix Prize or GE Flight Quest. Problems or needs are posted for any solver anywhere in the world to solve. Sometimes the problem can be simply stated (such as "prove Fermat's last theorem") but often they're accompanied by big datasets. For example, GE Flight Quest published data on scheduled and actual flight departure and arrival times as well as external data such as weather. A variety of techniques, ranging from 3D visualization to latent Dirichlet allocation, are then employed by solvers in an attempt to achieve the best result, typically but not always measured quantitatively.

Such innovation is typically accelerated. For example, the molecular structure of the Mason-Pfizer monkey virus retroviral protease hadn't been solved, even after 15 years of study by researchers using advanced computer models. It was solved in less than three weeks by contestants who were "ordinary people, not molecular biologists" via an open challenge conducted through an online, gamified site called Foldit.[8,9] Innovation via such mechanisms is also often of higher quality, due to simple math: there are more technical experts in any given area outside of a given company than within it. It can also often be less expensive, because a prize or licensing arrangement is typically only awarded for results, typically meeting specific thresholds, unlike employees or contractors, who are paid for effort, regardless of whether there are any practicable results. As Figure 4 shows, the cloud again plays a central role. As Jeff Weedman of P&G put it in describing the transformation from traditional innovation to the new model, "the new communication systems meant that we . . . could reach all the people that had [technical] capabilities."[10]

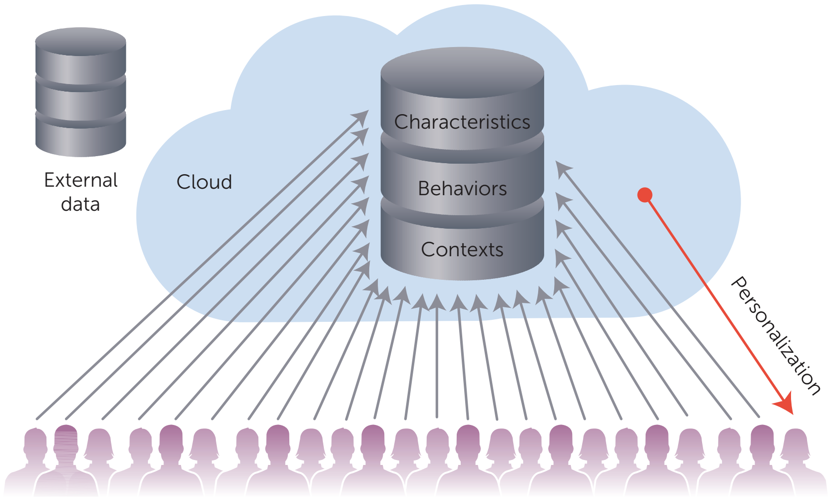


FIGURE 3. Collective intimacy high-level architecture. The cloud supports customer intimacy by aggregating data from customers to personalize and deliver services and recommendations.

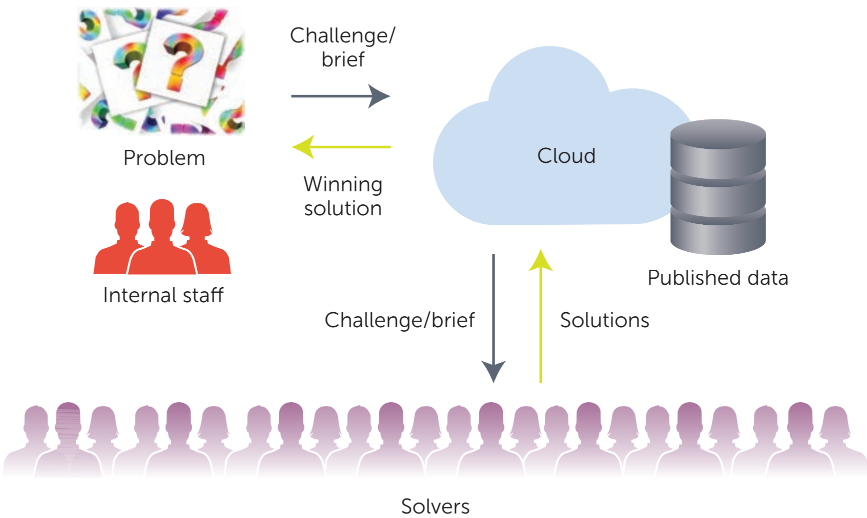


FIGURE 4. Accelerated innovation high-level architecture. The cloud allows firms to connect with experts outside of the organizational boundaries to produce novel solutions to problems.

In addition, the cloud and related information technologies can help accelerate innovation, for example, by offering speedy, cost-effective resources for conducting Monte Carlo simulations, enabling synchronous and asynchronous collaboration between researchers, or accelerating the digital portion of next-generation products through platform services and microservices or physical component prototyping through 3D printing.

Not Just Business, but Government

A companion piece in this issue of *IEEE Cloud Computing* delves into the role of cloud computing for governments.[11] The parallels are clear. For example, the California Natural Resources Agency leveraged cloud computing in an information excellence strategy to accelerate data acquisition and improve real-time response to rapidly changing natural disasters. Singapore has pursued solution leadership, connecting everything from driverless buggies to elderly patients to the cloud. Estonia pursued a customer (that is, citizen) engagement approach using individual cryptographic keys to better identify and thus engage with citizens across a broad range of government services. And governments can use the cloud to accelerate mission innovation.

THE CLOUD CAN PLAY A ROLE IN COST REDUCTION, PERFORMANCE OPTIMIZATION, FLEXIBILITY, AND USER EXPERIENCE. At its best though, it's the nexus of strategic competitive differentiation through the use of information excellence to optimize processes against any of a variety of goals such as cost, time, quality, or sustainability; solution leadership to tie physical endpoints to cloud-based services, social networks, and communities; collective intimacy to offer personalized services based on big data algorithms; and accelerated innovation through cloud-mediated contests, challenges, innovation networks, and idea markets.

References

1. *Unlocking the Cloud* research paper, Oxford Economics and SAP, 2013; http://scn.sap.com/docs/DOC-40516.
2. J. Weinman, *Cloudonomics: The Business Value of Cloud Computing*, John Wiley & Sons, 2012.
3. J. Weinman, *Digital Disciplines: Attaining Market Leadership via the Cloud, Big Data, Social, Mobile, and the Internet of Things*, John Wiley & Sons, 2015.
4. M. Treacy and F. Wiersema, *The Discipline of Market Leaders: Choose Your Customers, Narrow Your Focus, Dominate Your Market*, Basic Books, 1997.
5. S. Goodyear, "New York Wants Google Maps to Discourage Left Turns," *The Atlantic*, CityLab, 9 July 2015; www.citylab.com/commute/2015/07/new-york-wants-google-maps-to-discourage-left-turns/398114.
6. X. Amatriain, "Big & Personal: Data and Models behind Netflix Recommendations," *Proc. 2nd Int'l Workshop Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications*, 2013, pp. 1–6; xavier.amatriain.net/pubs/BigAndPersonal.pdf.
7. H. Chesbrough, *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Harvard Business Review Press, 2003.
8. D. Praetorius, "Gamers Decode AIDS Protein That Stumped Researchers for 15 Years in Just 3 Weeks (VIDEO)," Huffington Post, 19 Sept. 2011; www.huffingtonpost.com/2011/09/19/aids-protein-decoded-gamers_n_970113.html.
9. S. Cooper et al., "Predicting Protein Structures with a Multiplayer Online Game," *Nature*, vol. 466, 5 Aug. 2010, pp. 756–760; www.nature.com/articles/nature09304.epdf.
10. D. Matthews, "An Open Innovation Discussion with Jeff Weedman of P&G," *Canadian Business J.*, 10 Aug. 2010; www.cbj.ca/an_open_innovation_discussion_with_jeff_weedman_of_p_g.
11. K.L. Jackson, "IEEE GovCloud: Making a Difference for Global Governments," *IEEE Cloud Computing*, vol. 2, no. 4, 2015, pp. 54–59.

JOE WEINMAN *is a frequent keynoter and the author of* Cloudonomics *and* Digital Disciplines. *He also serves on the advisory boards of several technology companies. Weinman has a BS from Cornell University and an MS from the University of Wisconsin-Madison, both in computer science, and has completed executive education at the International Institute for Management Development in Lausanne. Contact him at joeweinman@gmail.com.*

This article originally appeared in IEEE Cloud Computing, *vol. 2, no. 4, 2015.*

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.

EDITOR MIKE HINCHEY
Lero—the Irish Software Research Centre;
mike.hinchey@lero.ie

The Ecology of Software Ecosystems

Tom Mens and Philippe Grosjean, University of Mons

Software ecosystems—collections of software projects developed and used by the same community—are extremely complex. Comparing them with biological ecosystems can yield new strategies for improving their effectiveness and resilience.

The discipline of ecology studies the interactions among living things in the context of their physical environment. The dynamics of these interactions are influenced by energy, nutrients, gas exchange, and temperature, as well as other organic and inorganic materials in the environment. For example, a wide variety of marine species in coral reefs are supported in a fragile environmental equilibrium; even small temperature fluctuations can have mortal consequences within such an ecosystem.

Selected CS articles and columns are also available for free at **http://ComputingNow.computer.org.**

Biological ecosystem dynamics are traditionally represented by a trophic web, reflecting the energy flow among producers, consumers, and the environment; shifts in species populations and available resources affect an ecosystem's delicate equilibrium (see Figure 1a). An ecosystem should be robust enough to maintain a stable equilibrium despite fluctuating dynamics to support its biological communities.

SOFTWARE ECOSYSTEM

In their book *Software Ecosystem*, David Messerschmitt and Clemens Szyperski define a software ecosystem as "a collection of software products that have some given degree of symbiotic relationships."[1] Mircea Lungu describes it as "a collection of software projects which are developed and evolve together in the same environment."[2] Well-known examples of software ecosystems include programming language archive networks, mobile app stores, and various distributions of the Linux operating system.

Given the sociotechnical nature of software development projects that involve software components and the

2376-113X/15/$31.00 © 2015 IEEE Published by the IEEE Computer Society

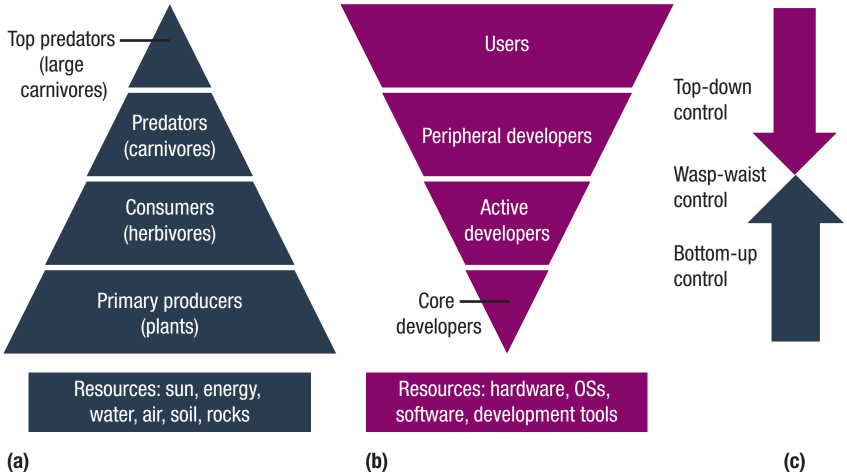


Figure 1. Schematic comparison of ecosystems. (a) Trophic web of a biological ecosystem. (b) A software ecosystem seen from a social viewpoint. (c) Ecosystems can be controlled from the bottom up—constrained by resources available to primary producers—or from the top down, driven by predators' consumption. In the wasp-waist control method, partial effects occur simultaneously from both directions.

project's contributors, including developers, testers, debuggers, and end users, software ecosystems can be compared to biological ecosystems in two ways:

- *Biological species ≈ software components*. From a technical viewpoint, software components can be compared to living species in a biological ecosystem. The software ecosystem contains all the hardware and software components required for developing, testing, deploying, and executing the software. As in a trophic web, some of these components (such as shared development platforms and software libraries) act as producers that are consumed by other software components.
- *Biological species ≈ project contributors*. In this social view, software project contributors are analogous to the living species in a biological ecosystem. As shown in Figure 1b, the project contributors form a trophic web consisting of core developers that produce the core architecture consumed by active developers, and so on. All these contributors share a common pool of project resources—such as version commits, issue reports, change requests, bug fixes, documentation, and tests.

CONTROL MECHANISMS GOVERNING ECOSYSTEM DYNAMICS

As Figure 1c illustrates, ecosystem dynamics can be controlled from the bottom up, constrained by the resources available to primary producers. Alternatively, they can be constrained from the top down, driven by predators' consumption. The wasp-waist control method combines both, with partial effects occurring simultaneously from both directions.

In the technical view of a software ecosystem, the bottom of the trophic web contains the set of components (including shared libraries) that constitute the core software architecture, upon which all other software components depend. In the social view, the trophic web contains a small number of core developers at the bottom, active developers higher up, peripheral developers even higher, and end users at the top, whose consumption creates the demand for more developers.

Some software ecosystems are primarily driven by input from core developers, with constraints related to limited resources such as budget, hardware, and personnel, or bottom-up controlled. Other software ecosystems are driven largely by change requests and bug reports coming from end users, or top-down controlled. Empirical studies aimed at understanding the relationship between control mechanisms and software ecosystem dynamics will lead to better software project management strategies.

ECOSYSTEM DIVERSITY

Resilience describes an ecosystem's ability to return to equilibrium once it's knocked out of balance. A biological ecosystem with high diversity is more likely to be resilient because some species compensate for others that disappear. Biodiversity could also play an important role in the success and resilience of software ecosystems, from both the technical and social points of view.

In the technical analogy, environmental perturbations can be caused by the loss of resources, either due to lack of interest or competing software components that become higher priority. They can also be caused by changes in the development process and the technology used, such as the introduction of new programming languages, new OS versions, and new development and version control tools. Accommodating a wide variety of software components could increase the software ecosystem's resilience or ability to adapt to change. This diversity could take many forms: the functionality offered, the programming language used, the supported OS, the end users targeted by the software, and so on.

This article originally appeared in Computer, *vol. 48, no. 10, 2015.*

In the social view, environmental disturbances can also lead to a loss of resources—for example, legacy projects might no longer be viable because the OS or programming language for which they were developed has become outdated. Increased diversity among the contributing developers—such as fluency in different programming languages and OSs or skill sets that include a range of activities including testing, debugging, documentation, translation, and so on—improves the ecosystem's resilience.

Many different metrics have been proposed to measure species diversity. Daryl Posnett and his colleagues relied on measures of relative entropy to study the activity focus of each software component (and the activity focus of each project contributor), and related this to the likelihood of introducing software defects into these components.[3] In a recent survey, Benoit Baudry and Martin Monperrus explored the different facets of software diversity, stressing the importance of identifying the underlying principles driving diversity's constant renewal.[4]

SURVIVAL OVER TIME

The theory of Darwinian evolution is generally used to describe the major mechanism driving *biological speciation*, which is when one species differentiates into two. According to this theory, all species compete in the same resource pool and survive by increasing their fitness or ability to reproduce. Over time, the most competitive genetic traits proliferate. This is akin to the evolution of software ecosystems, in which successful traits continue to proliferate and less successful traits disappear from future software versions.

In statistical ecology, the technique of survival analysis has been used to estimate, compare, and model the survival rates of animal populations in their ecosystems. Survival analysis is also gaining traction in software ecosystem research, determining the main factors that predict survival of software components or project contributors within their ecosystem.

FURTHER READING

For additional information on business software ecosystems and the usefulness of the evolution-based changes in software ecosystems, the authors recommend *Software Ecosystems: Analyzing and Managing Business Networks in the Software Industry* (Edward Elgar Publishing, 2013) and *Evolving Software Systems* (Springer, 2014).

Although research on software ecosystems is thriving, the complex dynamics of these ecosystems are still not well understood. Adapting biological theories and measures to the context of software ecosystems will help researchers come up with new strategies to improve the effectiveness and resilience of software ecosystems. ■

ACKNOWLEDGMENTS

This research was conducted as part of the Belgian ARC research project AUWB-12/17-UMONS-3, entitled "Ecological Studies of Open Source Software Ecosystems."

REFERENCES

1. D.G. Messerschmitt and C. Szyperski, *Software Ecosystem: Understanding an Indispensable Technology and Industry*, MIT Press, 2003.
2. M. Lungu, "Towards Reverse Engineering Software Ecosystems," *Proc. Int'l Conf. Software Maintenance* (ICSM 08), 2008, pp. 428–431.
3. D. Posnett et al., "Dual Ecological Measures of Focus in Software Development," *Proc. IEEE Int'l Conf. Software Eng.* (ICSE 03), 2003, pp. 452–461.
4. B. Baudry and M. Monperrus, "The Multiple Facets of Software Diversity: Recent Developments in Year 2000 and Beyond," to be published in *ACM Computing Surveys*, 2015.

TOM MENS is a professor at the COMPLEXYS Research Institute at the University of Mons. Contact him at tom.mens@umons.ac.be.

PHILIPPE GROSJEAN is a professor at the COMPLEXYS Research Institute at the University of Mons. Contact him at philippe.grosjean@umons.ac.be.

Subscribe today!

IEEE Computer Society's newest magazine tackles the emerging technology of cloud computing.

computer.org/cloudcomputing

IEEE computer society

IEEE COMMUNICATIONS SOCIETY

Premo: DreamWorks Animation's New Approach to Animation

Paul Carmen DiLorenzo
DreamWorks Animation

Creating an animated film takes several years, hundreds of highly skilled artists across many different departments (see the "CG Animated Film Production Process" sidebar), and advanced software tools. At DreamWorks Animation, our artists use a mixture of third-party and internally developed tools supported by a high-performance computing infrastructure to digitally create our animated films. We continually update our tools and take advantage of the latest hardware developments to meet our filmmakers' demands to produce more captivating characters and stunning visuals with each new movie.

In recent years, those hardware developments presented a new challenge. In the past, computer performance roughly doubled every two years, following Moore's law, which states that the number of transistors doubles every two years on a CPU. However, in the last decade, we have seen these additional transistors serve more to increase the number of CPU cores than increase the speed of individual cores. For the software industry as a whole, this has led to a crisis, and for the high-performance software we use, the problem is particularly acute because we rely on that increased core speed to deliver faster software to our artists.

To respond to this paradigm shift and scale with new hardware, we rearchitected our proprietary animation platform to take advantage of multiple-core computing. Premo is our new animation tool engineered by DreamWorks Animation technologists to enable artists to manipulate character performances in our movies.[1] Premo uses Apollo (www.dreamworksanimation.com/apollo), our new cloud-based digital design and CG media platform used for the creation and delivery of images, video, and other forms of media.

Premo was first used on the feature-length animated film *How To Train Your Dragon 2*. Since then, DreamWorks has made a full transition to Premo, and it has been used on the films *Home* and *Penguins of Madagascar* and is now being used on five films currently in production. Premo has exceeded our artists' expectations in both performance and usability and has resulted in greatly improving animator productivity and satisfaction.

Premo has not only resulted in a better user experience but also some key returns on investment for DreamWorks Animation. The training time for a new animator on Premo has been reduced to one to two weeks, compared with the 13-week course for the previous software. This enables animators to work on a movie soon after they are hired, whereas before we had to wait for them to be comfortable with proprietary tools. Initial studio estimates project up to a 20 percent reduction in production costs for movies using Apollo, starting with movies released in 2015. In addition, we feel that Premo offers audiences the highest quality entertainment experience with characters that provide the fullest range of visual emotion and expression.

Inspiration and Design

When designing Premo, we were able to leverage the many man-years of experience from our animators, allowing us to develop an application specifically suited to animation. To take advantage of this experience, we kicked off Deep Dive groups to focus on the main areas of the application. These groups discussed a specific topic for a few weeks and generated a substantial number of ideas. These ideas were distilled down into a set of Guiding Design Principles that outline key project requirements:

 2376-113X/15/$31.00 © 2015 IEEE

CG Animated Film Production Process

CG animated films are more complex than most people think. Each generally takes about five years to develop as a result of the planning and production of all the layers and assets that go into making a film. Great films begin with great concepts. Some of our ideas are completely original, whereas others are inspired by a variety of sources, including children's books and comic strips. Once we've settled on an idea, the first step is to write a script.

Storyboards

Once a script is ready, storyboard artists will imagine how the words will translate into actions and pictures by drawing a series of sketches to tell the story. The drawings are digitally strung together to create a story reel. (Imagine a flipbook that lets you see how the drawings flow together.) We combine that with temporary music, sound effects, and dialogue and work with the movie in this form for about 18 months.

Visual Development

Once the story reel is underway, our visual development department begins to plan the look of the film, developing the style, tone, color and overall artistic approach to each and every sequence. Everything has to be designed, from the major characters to the smallest of props. Thousands of drawings, paintings, blueprints, sculptures, and models later, our development artists have designed a fantasy world and characters to tell the story.

Modeling and Rigging

Modeling artists digitally sculpt the characters and environments in our films by collaborating with the art department to realize design concepts as tangible 3D forms. The modelers start with this wire frame sculpture that breaks down the design into workable geometry. A character technical director (CTD), also known as a *rigger*, will determine how the character must move and where the bones, muscle, and fat would be under their skin. Joints and various mathematical operations are then employed throughout the character's body, face, hair, and clothing to make it bend and deform like a living creature. Finally, rigging works with animation to design and build an extensive set of controls for the character so animation can pose every part of this digital puppet and bring it to life.

Layout

Layout artists interpret and recreate the hand-drawn 2D storyboard panels in a 3D CG environment. In doing so, these artists determine the 3D camera placement and motion. Working with rough versions of the characters, lighting, effects, and environments, artists create the cinematography for the film.

Character Animation

Once the sequence is working well in layout, the animators start bringing the characters to life in the computer. They articulate the thousands of controls that were created during the character-rigging phase to bring each character to life and to synchronize them to the voice performances.

Surfacing

Coming out of modeling, characters, props, and environments are flat and grey. The surfacing artists add the colors and textures to these elements, making surfaces look smooth and shiny like glass, bumpy and gritty like dirt, and fuzzy and soft like wool.

Effects

In a live-action film, it's easy to photograph things like leaves blowing in the wind, waves at the beach, or even footprints in the sand. In computer animation, these simple things are all designed and animated by the effects artists. In other words, if it's not acting, but it moves, it's an effect.

Lighting

Lighting artists utilize the computer to "paint" with light, bringing the final color, look, and illumination to the film. Lighting is the first time we get to see animation, surfaces, grass, trees, water, crowds, and effects all working together. Lighting does this by creating illumination for the scene. It creates the mood and atmosphere to support the story. Lighting leads the viewers' eye to the critical elements of the frame so that the audience is looking exactly when and where the filmmakers want them to look.

- direct control to "reach in and grab" the character;
- fast refresh of characters and environments in real time;
- representative high-resolution, fully deforming characters and environments, indicative of the final movie;
- creative workflows that allow animators to pose, draw, and explore in a more natural way;
- fluid and seamless flow between different tasks such as posing and adjusting motion across time;
- flexible and fully configurable interfaces that allow multiple views of animation controls;
- intuitive browsing, previewing, and editing of the entire movie segmented into "shots"; and
- immersive experience that keeps the animator "in the moment" with their creative thoughts and ideas.

To create a clear and consistent vision for the project, we borrowed the storyboarding approach

© 2015 DreamWorks Animation

Figure 1. Premo digital pen interface. An artist is using a Wacom Cintiq and a digital pen in Premo to edit Hiccup, the main protagonist character of the *How to Train Your Dragon* franchise.

© 2015 DreamWorks Animation

Figure 2. Premo character complexity. The model for the Bewilderbeast, a new dragon character introduced in *How to Train Your Dragon 2,* is 20 times the size of Toothless, the featured dragon from the original *How to Train Your Dragon* movie.

that our filmmakers use by creating mockups. Many different mockup types were used, such as images, movies, and Flash animations. These mockups helped demonstrate the interaction and specific workflows desired by animators inside of Premo.

Transition to Premo

Prior to Premo, animators used Emo, DreamWorks Animation's previous generation proprietary software, which was used to animate character performances in more than 25 of the studio's animated features. When developed, it was a state-of-the-art, award-winning tool; however, its architecture did not adequately support the multicore computing model that was needed to meet the interactivity goals of the project.

With Emo, changes made to character movements were typed into a spreadsheet-like interface, and it required significant time to calculate and render the results on screen. To help speed up this compute process, low-resolution versions of the characters were used, and environmental detail was removed. However, this meant animators could no longer experience their work as it would appear in the final movie, forcing them to make many guesses in animation performance, which often later required additional fixes. This method was not only time-consuming, but it also limited the number of iterations and broke the stream of creativity.

Premo's rearchitecture, based on modern multicore computing, has resulted in greatly enhanced real-time performance and interactivity. As a result, with the transition to using Premo, the animators' workflows and experiences radically changed, providing a variety of new capabilities and functionality. To fully capitalize on the Premo rearchitecture, we deployed to each animator a HP z820 workstation that uses Intel IvyBridge with 10 cores at 3 GHz, 96 Gbytes of RAM, Nvidia K5000 video cards with 4 Gbytes of RAM, and a 960-Gbyte solid-state drive (SSD).

Animator Workflows and Experience

The greatest experience shift for animators is to no longer need to wait for the recalculation of their scene to see the result of edits. They can quickly iterate on their animation, try different ideas, add more subtle motion and expressions, and ultimately realize their vision for the character performance.

This significant speed improvement enabled new workflows to be introduced. For example, animators can now work interactively with a pen on a pressure-sensitive tablet. They can precisely control a characters' face and body by simply moving their pen to the exact degree desired and receive real-time feedback (see Figure 1).

Animators desire the ability to work with the highest quality representation of the characters and environments. With Premo, animators manipulate high-resolution, fully deforming characters in their scene. They can load many characters and detailed environments to judge the character's performance in the context of the scene. Figure 2 shows a loaded scene in Premo where two Bewilderbeast dragons are battling. The Bewilderbeast dragons are large, detailed characters, yet Premo is able to stay interactive even with this level of complexity.

Advanced Animation Features

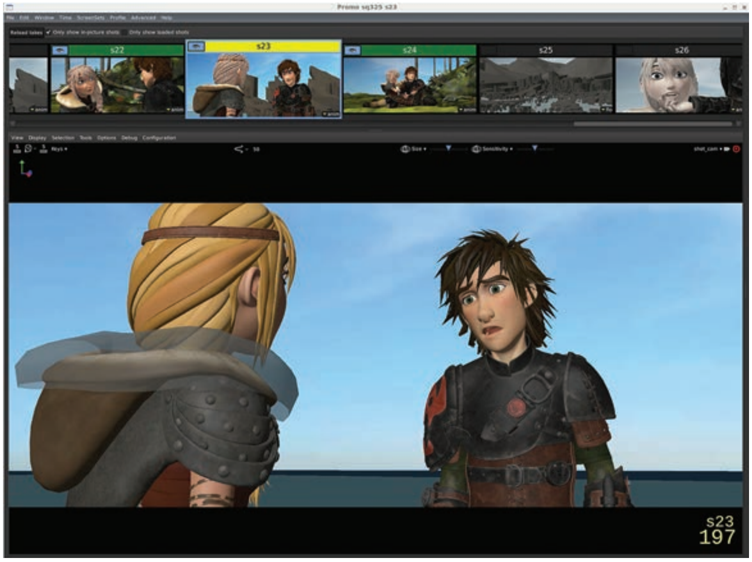
The guiding principle of creativity was a high priority for the animators. Toward that goal, we added a fast, high-quality drawing system into Premo that provides the ability to sketch in a shot using a digital pen (see Figure 1). Many of our highly trained 2D animators, whose natural language includes being able to draw by hand, use the drawing tools to develop ideas for their work. Animation supervisors use these drawings to provide notes to the animation team. This type of collaboration has proven to be faster in communicating character performances.

To help animators animate across time, we added a motion path and ghosting feature into Premo (see Figure 3). These features allow the animators to see the character in past and future frames, enabling them to effectively work across multiple frames in a single view. The motion paths provide a simple curve visualization and editing capability for a single animation control, and the ghosting feature provides a full character representation over time.

Continuity of character performances across multiple shots is important to maintaining visual consistency. Previously, our animators were confined to working within a single shot and were required to close and reload the application to work on other shots. To evaluate continuity between shots, they would have to generate movies for each of the shots, which was tedious and time-consuming. In Premo, we created the ability to open an entire sequence, choose the shots to be worked on, and add them to the session (see Figure 4). The animators can then make edits on any of the active shots and see how those changes play with the surrounding shots in the sequence. Because Premo is connected to Apollo's cloud platform, it can access and play videos created by other departments in the production pipeline. For example, they can access the latest renders from the lighting department of the shot they are currently working on to easily compare versions.

© 2015 DreamWorks Animation

Figure 3. Premo motion path and ghosting feature. With motion paths (purple line) and ghosts of the character over time, the animator can visually and quickly set the precise position and timing they want for the character's motion.



© 2015 DreamWorks Animation

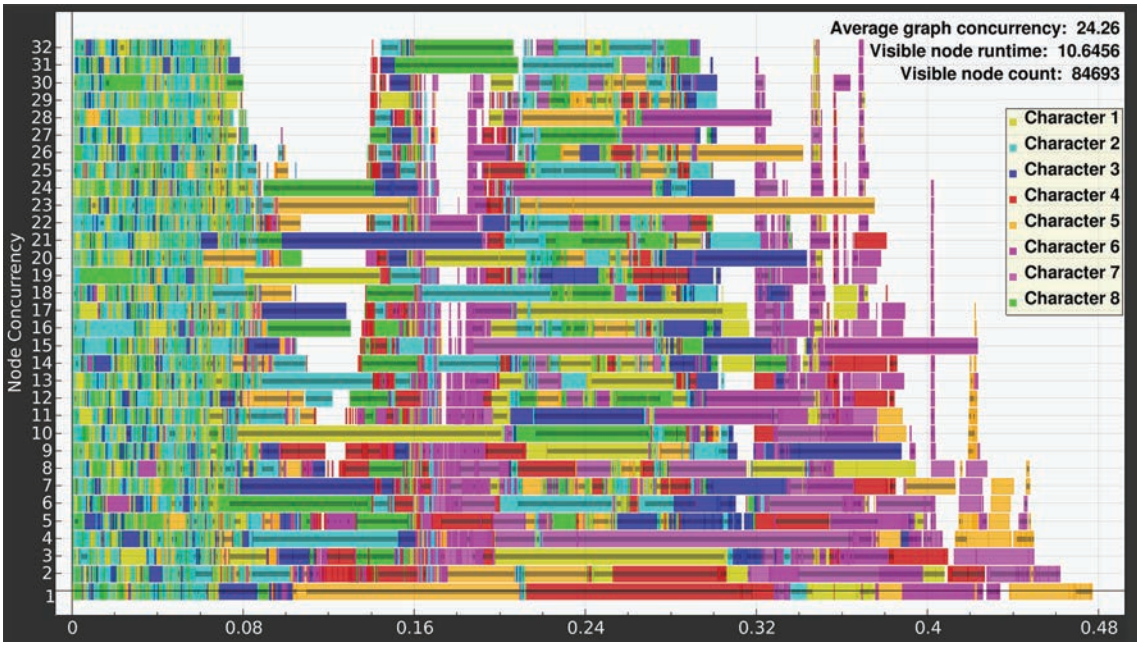
Figure 4. Premo Shot Browser. Using the Shot Browser viewer at the top of the Premo interface, animators can quickly switch between shots, choose the department view (such as layout or lighting) they want to access, hover over a shot to see a movie clip, or choose a set of shots to play continuously in the Scene View (main view) to ensure consistency in their animation.

Architecture and Technology

Three architectural goals helped us achieve the features and animator experience of Premo. The first is to maximize utilization of computational resources by distributing heavy computations across local cores and continue to be highly scalable as hardware capabilities increase.

Even with this highly scalable architecture, the application may not be able to maintain interactive rates given a sufficiently heavy workload. Therefore, the second goal is to build an architecture that treats computation as a service. This service must completely decouple evaluation from the client to ensure that the application is consistently responsive to users.

By separating the computation from the client, we are no longer restricted to doing computation on the local cores. This enables the third goal, which is the ability to allocate more cores to the artist when required. Because we expect the complexity of characters and environments to continually increase in order to achieve our filmmakers' goals, we need to be able to bring additional computing



© 2015 DreamWorks Animation

Figure 5. LibEE visualization tool. LibEE shows the execution of each node and stacks them when they are executed at the same time to show the degree of parallelism. Colored areas mean that the cores are being used, and the white space shows inactivity. This example depicts a single evaluation frame of eight characters on a 32-core machine. During the evaluation, there are times when all 32 cores are busy, indicating that performance is limited by the hardware resources, rather than the rig itself or an evaluation engine. Because future machines will almost certainly have more cores, this means we are well placed to continue to scale performance of our engine with future hardware.

resources to bear if the local animator workstation is insufficient.

These goals ensure that Premo can achieve the Guiding Design Principles even as workload complexity increases over time.

LibEE

DreamWorks Animation partnered with Intel to develop LibEE,[2] a high-performance, multithreaded dependency graph evaluation engine inside of Premo. LibEE is designed to evaluate the characters in a shot as efficiently as possible.

Before animators can create character performances, our character rigging department builds the skeletal structure of a digital puppet, called a *rig*. The rig defines the available controls on the character that the animator can use as well as the degree of motion that can be applied to those controls. Some of our characters have as many as 5,000 controls, and a rig can contain up to 150,000 individual *nodes*, each of which is an element of computation that performs one particular action within the rig. All these nodes are chained together into what is known as a *dependency graph*.

LibEE uses Intel's Threading Building Blocks library (http://threadingbuildingblocks.org) to achieve fast, scalable evaluation. The primary benefit of LibEE over other dependency graph evaluation engines is its ability to evaluate multiple parts of the character in parallel (such as arms, legs, and other independent parts of the character), providing a significant speedup over more traditional nonthreaded evaluation systems, up to 10 times in some cases.

One important concept is that in order for our rigs to evaluate in parallel, not only does the evaluation engine need to be multithreaded, but the rigs themselves also need to be structured to allow graph parallelism. As a simple example, if a rig contained a set of nodes that each attached one toe to a foot, and those nodes were connected in a series that each depended on the one before, then the evaluation engine would have to evaluate each toe at a time. However, when the nodes are connected in parallel, so there is no dependency between individual nodes, then all the toes can be evaluated at the same time.

Building a rig that allows the engine to extract maximum scalability from the workload was a new challenge for our production artists.[3] To help meet this challenge, we implemented profiling tools that allow riggers to visualize the data flow and node evaluation in a graph (see Figure 5). Artists can identify which character components are running in parallel, determine occurrences of graph serialization bottlenecks, and discover where unexpected dependencies may exist between parts of the rig. The tool highlights nodes on the critical path (along

© 2015 DreamWorks Animation

Figure 6. Premonition system process. (a) When an animator edits a character, that prompts Premonition to begin recomputing the surrounding animation frames, indicated by the orange bar. (b) The brown bar clears as those adjacent frames compute. (c) The animator can then begin playback to watch the animation, which can display in real time because Premonition has precomputed these results. At this time, Premonition will continue to process the remaining frames in the background. (d) Finally, the animator can change time manually, and Premonition will adapt and predict new frames to compute based on the new time.

the bottom horizontal bar), which determine the overall best possible runtime of the graph.

Premonition

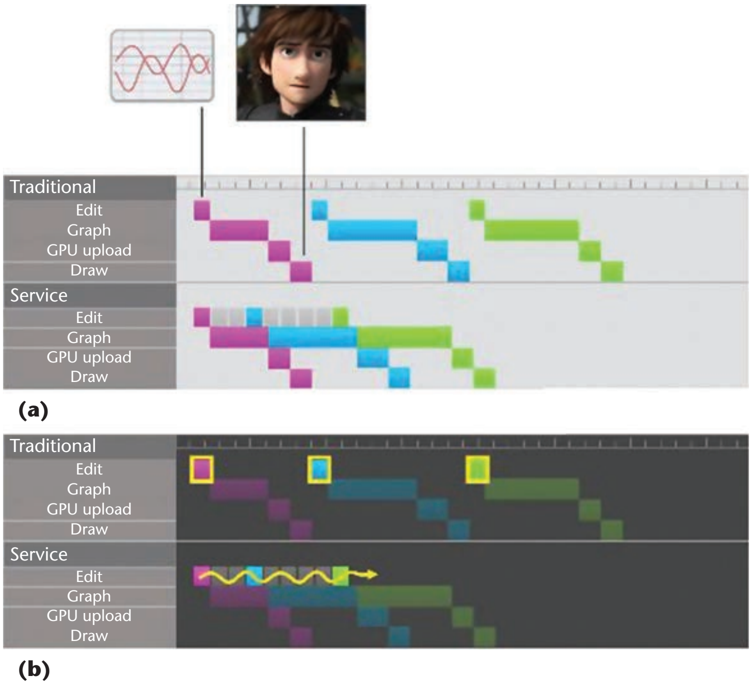
Premonition is one of the key systems inside of Premo. Its purpose is to take advantage of the idle time (sometimes fractions of a second) between the edits an animator will make on a character. While animators are posing on a single frame, they are more interested in seeing how that edit affects the motion of the character over time. Therefore, after an animator makes an edit, Premonition will start to compute adjacent frames immediately (see Figures 6a and 6b). By taking advantage of this animator workflow, we maximize utilization of the computational resources on the workstation.

Because we tuned our architecture toward animator workflows, Premo allows the animator to make an edit and playback the shot while Premonition processes the shot in the background (see Figure 6c). Premonition also has the ability to adapt and predict (see Figure 6d). As the user changes to a different time in the shot, Premonition will adapt and adjust the frames it computes based on the new time the user wants to view. From there, it will try to predict and compute the frames the user will view next.

High-Performance Evaluation Architecture

Premo's high-performance evaluation architecture uses a layered architecture with LibEE at the bottom and the client layer on top. The client layer contains the UI providing user input in the form of animation curves and output in the form of geometry to display on screen. Between LibEE and the client (the middle layer) is the Graph System, which acts as the application's interface into the graph.

To treat computation as a service, the Graph System manages evaluation requests through a subscription model, where the client subscribes to a



© 2015 DreamWorks Animation

Figure 7. Treating computation as a service in Premo. (a) The traditional approach couples the application with the graph. This results in the application waiting for the graph results and under-utilization of the graph. In comparison, our service-based approach allows the application to stay responsive and fully utilizes the graph. (b) With the traditional approach, the user would receive results only when an edit is fully realized by the system. This leads to a choppy experience for the animators. In contrast, our service-based approach produces a continuous, analog experience for the animators. (Courtesy of DreamWorks Animation.)

particular output of interest from the Graph System over specific time frames. The Graph System uses these subscriptions to orchestrate graph evaluation with LibEE and then batches these requests for optimal parallelism. When LibEE finishes an evaluation, the Graph System pushes the results back to the client to immediately display the results.

To illustrate the unique benefits of treating computation as a service, let us step through a single cycle of an application utilizing the traditional approach (see the top row of Figure 7a). This approach begins when a user edits animation curves. This kicks off a graph evaluation to process the curves and compute new character geometry. Finally, the geometry uploads to the workstation's GPU to display the results on the screen.

In the traditional approach, graph evaluation is coupled with the application event loop where the main UI thread is constantly waiting for the graph to finish before it can update. In our service-based approach (see the bottom row of Figure 7a), we have essentially compressed all the colored blocks to the left, creating a full pipeline of edits, evaluation, and display. By decoupling each of these tasks, user edits can be processed at a higher rate than graph evaluation, which is represented by the gray boxes. Furthermore, computing resources are now fully utilized by continuously keeping the graph busy, resulting in an increase in overall throughput of our pipeline.

The outcome of this approach for the animator experience is significant. As seen in Figure 7b, the service architecture inside Premo returns 3D animation back to a continuous and analog experience, not a discrete one as experienced in traditional 3D editing applications. Creative processes, such as drawing on paper, are inherently analog, smooth, and fluid. When the industry moved to computer animation, the creative process became discrete due to the nature of engineering. However, with Premo eliminating the wait time on the graph, animators are able to return to an analog experience, and the result is a natural interface for 3D animation.

Solving Advanced Animation Challenges

Character constraints are created when an animator wants one character to drive the motion of another. For example, if the character Hiccup is riding his dragon Toothless, an animator can constrain Hiccup to Toothless in Premo. When the animator moves Toothless, Premo will automatically compute the new position of Hiccup. A more complex example would be characters locking arms and influencing each other's motion. The challenge is that character constraints often introduce infinite evaluation loop cycles in our graph. Furthermore, when adding numerous constraints, it is difficult to maintain interactive rates. Therefore, we build in native support for handling cyclic dependencies. In the Graph System, we manage constraints and track the resulting cyclic dependencies, while leaving LibEE as a directed acyclic graph (DAG). To evaluate the cyclic dependencies, we do a multiple-pass planning stage based on constraint priorities. The resulting plan is then executed by LibEE.

The second animation challenge is the seamless integration of editing and simulation. Character simulations (such as hair and clothing) increase the visual fidelity we give to the animators. We want the animator to be able to see live simulation results for the frame they are viewing when they are posing a character. When an animator manipulates a character that has simulations, the simulation requires the evaluation of previous frames before the current frame can be displayed. This requires extensive computational resources. In contrast, when an animator manipulates a character without simulations, the dependency on previous frames does not exist and the application only needs to evaluate the current frame. The challenge is to ensure the architecture is aware of these temporal requirements and can do the necessary computation to achieve the desired result for the animator. To solve this challenge, we introduced the concept of time-based dependencies into the architecture. Simulations are required to describe their time-based dependencies to ensure that the current frame being displayed to the animator is always up to date.

From the onset of the project, we set out to build a tool for animators, inspired by animators, with the motto of "no compromises." The result has been the achievement of both the design and architectural goals we established for ourselves. We know our tools will continue to make effective use of future hardware, enabling us to keep up with our demanding production environment.

Acknowledgments

I thank Martin Watt, Matthew Gong, Fredrik Nilsson, Anna Lee, and Bill Ballew for their help and guidance with this article. I also thank all the engineers, artists, and managers who worked on this amazing software (past and present) and for creating significant IP in this area as well as the studio leadership for their support of this program. Special thanks also to Theresa-Marie Rhyne for encouraging the development of this article.

References

1. M. Gong et al., "Premo: A Natural-Interaction Animation Platform," *ACM SIGGRAPH 2014 Talks*, 2014, article no. 3.
2. M. Watt et al., "LibEE: A Multithreaded Dependency Graph for Character Animation," *Proc. Digital Production Symp.* (DigiPro), 2012, pp. 59–66.
3. G. Zimmermann, K. Ochs, and R. Helms, "Building Highly Parallel Character Rigs," *ACM SIGGRAPH 2014 Talks*, 2014, article no. 27.

Paul Carmen DiLorenzo *is the R&D project manager for the Animation and Rigging Program at DreamWorks Animation. Contact him at paul.dilorenzo@dreamworks.com.*

Contact department editor Mike Potel at potel@wildcrest.com.

This article originally appeared in IEEE Computer Graphics & Applications, *vol. 35, no. 4, 2015.*

Daniel E. Geer Jr.
In-Q-Tel

Children of the Magenta

The term "children of the magenta" traces to 1997, when American Airlines captain Warren Vanderburgh said the industry has made pilots too dependent on monitoring the magenta lines on the machines that are really flying the plane (http://99percentinvisible.org/episode/children-of-the-magenta-automation-paradox-pt-1).

William Langewiesche's article analyzing the June 2009 crash of Air France flight 447 comes to this conclusion: "We are locked into a spiral in which poor human performance begets automation, which worsens human performance, which begets increasing automation" (www.vanityfair.com/news/business/2014/10/air-france-flight-447-crash).

University of Miami professor Earl Wiener proposed a set of "laws" that include every device creates its own opportunity for human error; exotic devices create exotic problems; and digital devices tune out small errors while creating opportunities for large errors.

Langewiesche's rewording of these laws is that "the effect of automation is to reduce the cockpit workload when the workload is low and to increase it when the workload is high" and that "once you put pilots on automation, their manual abilities degrade and their flight-path awareness is dulled: flying becomes a monitoring task, an abstraction on a screen, a mind-numbing wait for the next hotel."

Nadine Sarter of University of Michigan said that such "de-skilling is particularly acute among long-haul pilots with high seniority." As Langewiesche added, "Beyond the degradation of basic skills of people who may once have been competent pilots, the fourth-generation jets have enabled people who probably never had the skills to begin with and should not have been in the cockpit."

The situation in aviation is precisely the situation we are in with cybersecurity. Human error is rampant at all levels. There is a cacophony of calls for cybersecurity automation. The most experienced people are no longer directly solving problems hour after hour but rather superintending largely automated processes. More and more, digital devices tune out small failures, whether they be attacks, misconfigurations, version mismatches, or service disconnects. Like airplanes automated enough that anyone can fly them, anyone can ostensibly operate the digital devices that are unarguably society's predominant risk vector. Therefore, there's a guarantee of large errors at some future point—errors that no one still in practice will handle. When successful automation makes particular threats increasingly unlikely to appear, the interval between failure events grows longer. As the latency between failure events grows, the assumption that safety has been achieved also grows, fueling increased dependence on what is now a positive feedback loop (http://geer.tinho.net/geer.sfi.2x14.txt).

Vanderburgh's "children of the magenta" also applies to cybersecurity in another way: you shouldn't run a cybersecurity detection and response operation via on-the-fly reprogramming of our equivalent of the Flight Management Computer. In 2013, *Aviation Week* editorialized that "there needs to be a new performance-based model that requires flight crews to log a minimum number of hand-flown takeoffs and departures, approaches and landings every six months, including some without autothrottles. Honing basic pilot skills is more critical to improving airline safety than virtually any other human factor" (http://aviationweek.com/commercial-aviation/editorial-how-end-automation-dependency).

If you aren't regularly flying your cybersecurity airframe manually, you can and will become automation dependent. Then, just as with airplane pilots, in a rapidly changing environment, you'll lose situational awareness due to task saturation brought on by the automation itself. We can't allow ourselves to be so automation dependent that we can't turn off the automation and fly the plane. ■

Daniel E. Geer Jr. is the chief information security officer of In-Q-Tel. Contact him at dan@geer.org.

This article originally appeared in IEEE Security & Privacy, *vol. 13, no. 5, 2015.*

2376-113X/15/$31.00 © 2015 IEEE

Editors: James J. Hack, jhack@ornl.gov | Michael E. Papka, papka@anl.gov

Putting Regional Climate Prediction in Reach

Laura Wolf and Jim Collins | Argonne National Laboratory

Through a diversity of user facilities that include both major supercomputing centers and strategically situated sensing observatories, the US Department of Energy (DOE) is advancing the foundational science needed to inform climate change mitigation and adaptation strategies for policymakers, energy markets, and the public.

Fighting climate change is now a core component of the DOE's work. The agency's policy recommendations are informed in part by decades of data collected by a worldwide network of advanced observations and instruments. The DOE also supports numerous global and regional climate-modeling efforts and a robust program of computational climate science research aimed at obtaining more predictive, reliable, and trusted simulations of highly complex climate systems to help in decision making.

Regional climate modeling has a high potential for translating data into meaningful policy. One research team is using Mira, a 10-Pflops IBM Blue Gene/Q system at the Argonne Leadership Computing Facility (ALCF), to evaluate regional climate models (RCMs) against observational data—an important step in building confidence in the use of these models for projection.

Climate models comprise systems of differential equations that describe how a moving fluid's physical factors, such as velocity, pressure, and temperature, are related. The differential equations are the basis for computer programs that simulate the atmosphere or ocean—the key components of global climate models (GCMs), along with sea-ice and land-surface components. To extract useful information from a GCM, the model must simulate time spans on the order of many decades, frequently with the goal of projecting Earth's climate state one century out.

Climate models simulate the relevant physical processes affecting interactions among the atmosphere, oceans, and biosphere. Scenarios of past and future emissions of the

2376-113X/15/$31.00 © 2015 IEEE

> Most researchers make use of one of several community codes, some of which have been under development for decades.

important radiatively active gases and particles (as well as volcanic, solar, and other natural "forcings" that cause the climate system to change) drive these models.

Most researchers make use of one of several community codes, some of which have been under development for decades. These codes are written in accessible programming languages such as Fortran, C, or Python, typically using MPI and OpenMP to implement parallel constructs and the associated communication. Communities of users form around these models based on their ease of use, capabilities, and availability of reusable libraries. Thus, if a researcher wants to model something in particular and it's published and available to use, she needs only to process the model's output.

Climate science investigators often build several smaller component models and then couple them with the community code, where the interaction of these modeling components helps to provide a better understanding of the climate processes being investigated. This work leads to refinements of the component model as well as the community modeling framework, which often makes the code much more complicated and expensive to run.

Using massively parallel supercomputers at DOE Office of Science User Facilities, such as Mira at the ALCF, and Titan, the 27-Pflops Cray XK7 system at the Oak Ridge Leadership Computing Facility, a research team can run many scenarios concurrently at unprecedented spatial and temporal resolution.

With the huge number of nodes and cores on these machines, multiphysics and multiscale simulations considerably shorten the time to solution. In addition, the volume of model-generated output is massive, which in turn demands massive amounts of storage that are either available directly on or near these supercomputing systems.

The data output from GCMs, which typically operate at a spatial grid resolution ranging from 100 km to 300 km, is extremely useful for examining climate trends on oceanic and continental scales. But understanding how climate change can impact weather-sensitive businesses, such as agriculture, or the water budget in areas the size of a single state (or cluster of states) requires an RCM capable of resolving geographic areas in the range of tens of kilometers.

From Global to Regional

Generating climate projections at a regional level requires some method of downscaling—a technique that takes information known at large scales to make predictions at local scales. In climate modeling, downscaling generally is either based on an RCM (dynamical downscaling) or is empirically based on the relationship between climate model output and long-term observations (statistical downscaling).

Dynamically downscaled models are more computationally expensive than statistical models, but they're better at reducing the uncertainty of inputs, or parameters, such as the processes representing clouds.

Argonne National Laboratory climate and atmospheric scientist V. Rao Kotamarthi and his team are using Mira to investigate the effectiveness of dynamically downscaled climate models for predictive purposes. They used the Weather Research and Forecasting Model (WRF, pronounced "wharf"), a numerical weather prediction system, to generate more than 100 years of model simulations at regional scales with a spatial grid resolution of 12 km over the North American continent. They then assessed the performance of RCM data output for one variable—precipitation—to see how well it captured space and time relationships.

Kotamarthi's regional model contained roughly 10 million grid cells, and the team ran up to 10 jobs at a time on Mira. Approximately 1 million core-hours were needed for a one-year simulation on 512 nodes (or 512 × 16 cores). In this configuration, a 100-year simulation could be performed in approximately 50 days.

Temporal resolution involves the size of the time steps used in climate models. Kotamarthi's RCM used simulated time steps of 40 seconds and saved the output every three simulated hours—a timescale at which the team had the ability to explore climate changes that occur on a diurnal scale, such as thunderstorms or urban heat islands.

The researchers ran simulations using different boundary conditions for the time periods of 1980 to 2010 and 1995 to 2005 and then compared these simulations with observational data to validate calculation accuracy. After building confidence in their RCM, the research team carried out more simulations to predict how regional precipitation patterns would change in future decades.

The team found that downscaling added significant value to the simulations at the regional scale. The correlations in the RCM output showed similar patterns to the observational data and exhibited much better agreement with the observations than the global model did in capturing small-scale spatial variations of precipitation—especially over mountainous regions and coastal areas. Moreover, even when the simulations were averaged to the size of global models, the higher-resolution RCM retained more information and matched the observations better than models operating at 100- and 250-km resolution (see Figure 1).

Researchers anticipate that with the new generation of computing hardware that the DOE is acquiring, operating the climate models at higher spatial resolutions will improve the models' projection performance with observations at all spatial scales. ■

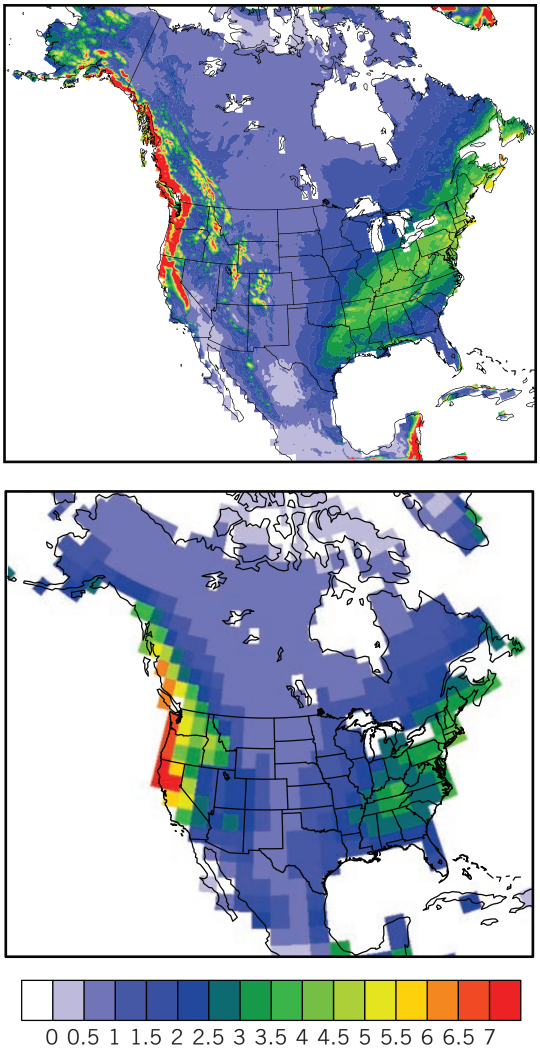


Figure 1. Average winter precipitation rate (mm per day) for a 10-year period (1995 to 2004) as simulated by a regional climate model with 12-km spatial grid resolution (top) and a global climate model with 250-km spatial grid resolution (bottom).

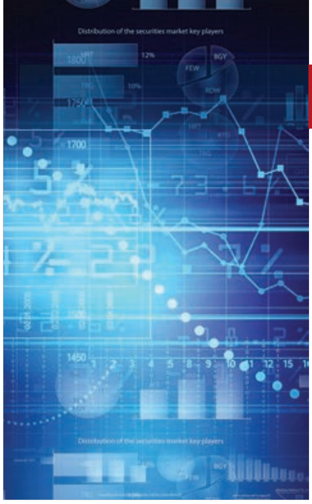
Acknowledgments

This research is supported by the US Department of Energy's Office of Science's Advanced Scientific Computing Research Program and under a military interdepartmental purchase request from the US Department of Defense's Strategic Environmental Research and Development Program (SERDP), RC-2242, through US DOE contract DC-AC02-06CH11357. The Argonne Leadership Computing Facility is a DOE Office of Science User Facility supported under contract DE-AC02-06CH11357. The Oak Ridge Leadership Computing Facility is a DOE Office of Science User Facility supported under contract DE-AC05-00OR22725. We gratefully acknowledge additional editorial assistance by Gail Pieper.

Laura Wolf is a science writer and editor for Argonne National Laboratory. Her interests include science communication, supercomputing, and new media art. Wolf has an MA in journalism from Columbia College Chicago. Contact her at lwolf@anl.gov.

Jim Collins is a writer and editor at Argonne National Laboratory. In his six years at the laboratory, he has provided communications support for diverse research areas including transportation, renewable energy technologies, and supercomputing. Collins has an MS in journalism from Northwestern University. Contact him at jcollins@anl.gov.

This article originally appeared in Computing in Science & Engineering, *vol. 17, no. 5, 2015.*



Editor: **V.S. Subrahmanian,** University of Maryland, vs@cs.umd.edu

Saving Rhinos with Predictive Analytics

Noseong Park, Edoardo Serra, and V.S. Subrahmanian, *University of Maryland*

The last author was astonished, when, at a computation and conservation meeting in June 2014, a pioneer in AI was heard proposing algorithms to patrol a game park to protect endangered species without any data on either animal locations or poaching incidents. It would seem self-evident that you cannot protect, say, a rhino if you don't know where it is.

In 2013, the Olifants West Nature Reserve in South Africa provided our team with 18 months of poaching incident data, as well as data about the movement patterns of a few rhinos. (Thanks to Stefan Bosman, Kirsty Brebner, Amy Clark, Tom Snitch, and Craig Spencer for facilitating access to this data.) More than 1,200 rhinos are poached annually in South Africa alone, as well as 27 in India's Kaziranga National Park and 24 in Namibia[1]; thus, protecting rhinos is a major priority for global wildlife conservation efforts. With a black-market price of US$300,000 per rhino horn,[2] poachers have every incentive to carry on their grisly trade.

Our team has been building a system called the Anti-Poaching Engine (APE), which has been transitioned to Kaziranga.[3] APE uses many technologies, such as drones—but a drone is nothing more than a dumb flying object. APE proposes methods to compute, on day d, a set of coordinated drone flight paths and on-ground ranger patrols for the next day ($d + 1$), so that as many rhinos as possible are protected. These novel planning algorithms build on sophisticated behavior models of both rhino behavior and poacher behavior—marking, to our knowledge, the first time ever that route planning has been fed real-world, data-driven behavior models as input.

Features

We took the Olifants West Nature Reserve and divided it into a total of 1,014 (400 m × 400 m) cells, covering the entire park (over 162 million square meters). We then came up with a set of features for each cell c that we would use in building our behavioral model:

- *DW1(c)*, *DW2(c)*, and *DW3(c)*, respectively, denote the distance of c from the nearest, second nearest, and third nearest sources of water. These features capture the fact that rhinos need water.
- *DR1(c)*, *DR2(c)*, and *DR3(c)*, respectively, capture the distance of c from the nearest, second nearest, and third nearest roads. These features were intended to capture avoidance (of human activity) by rhinos.
- *DH1(c)*, *DH2(c)*, and *DH3(c)*, respectively, denote the distance of c from the nearest, second nearest, and third nearest houses or buildings—also designed to capture avoidance of humans by rhinos.
- *DV1(c)*, *DV2(c)*, and *DV3(c)*, respectively, denote the distance of c from the nearest, second nearest, and third nearest sources of vegetation. These features capture the fact that rhinos need food.
- *MAXELEV(c)* and *AVGELEV(c)* show a cell's maximal and average elevation, obtained by randomly sampling 100 points in the cell.
- *STEEPNESS(c)* is the difference between a cell's maximal and minimal elevation, captured from the 100 randomly sampled points. This captures the fact that rhinos might not like to traverse steep slopes.
- *VISIT(c, i)* captures the number of rhinos visiting a cell at distance i from cell c (where $i = 1, 2$).

The elevation-related features served as a proxy for a rhino's inclination to travel steep paths. All these variables are also inextricably linked to poacher behavior; however, poachers could target animals other than rhinos. Using this base set of features, we set out to discover models of rhino behavior and poacher behavior.

Published by the IEEE Computer Society

2376-113X/15/$31.00 © 2015 IEEE

Poacher Behavior

A common statement about poacher behavior is that poachers like to target animals near roads (allegedly because the former are lazy). Our data showed that this is incorrect—there are much better explanations. Using association rule mining,[4,5] we obtained good predictors of poacher behavior. We found 75 association rules that had support of 2 percent or more, confidence of over 70 percent, negative confidence below 5 percent, and lift of 10 or more.

Given a condition *C* predicting that poachers will attack cell *c*, negative confidence is the conditional probability $\mathbf{P}(\text{Attack}(c)|\text{not } C)$. When confidence is high and negative confidence is low, this suggests that condition *C* can be used as a predictive beacon. When it is on, we can predict that the cell will be targeted by poachers with high probability, whereas when it is off, attacks occur with low probability. Interestingly, upon further scrutiny, most of the 75 rules boiled down to one single rule—cell *c* is likely to be targeted by poachers when

- *c*'s average elevation is less than 394 meters, and
- the distance from *c* to the nearest water source is 0 (that is, there is a water source in cell *c*).

In this case, support is 2.17 percent, confidence is 70.97 percent, negative confidence is 4.78 percent, and lift is 10.43. For this rule, lift is the ratio of the conditional probability of a cell being targeted by poachers given that the two conditions about the cell given earlier are true, to the overall unconditional probability of the cell being targeted by poachers. The lift of 10.43 says that cells that satisfy these two conditions are more than 10 times more likely to be targeted by poachers compared to random cells.

Another interesting rule about poacher behavior yields the following rule with a lower support. Cell *c* is likely to be targeted by poachers when

- *c*'s average elevation is less than 365 meters,
- there are at least two animals per day in a cell at most one hop away from *c*, and
- the distance to the nearest source of vegetation is fewer than two cells (that is, less than 800 meters).

Here, support is 1.08 percent, confidence is 78.57 percent, negative confidence is 5.8 percent, and lift is 11.55. When compared to the preceding rule, this rule has higher confidence and lift, but only half the support.

These rules show how predictive analytics can help shape knowledge in other fields. Although some conservationists and park rangers intuited that distance to roads was a good rule of thumb, it is not a good predictor of which cells will be poached—nor is it validated by the data. In addition to learning association rules that neatly differentiate between cells targeted by poachers and those that are not, we used other predictive models to accurately predict which cells will be poached, achieving an overall classification accuracy of 91 percent. However, we note that in our best classifier (we tested support vector machines, *k*-nearest neighbors, decision trees, AdaBoost, and restricted Boltzmann machines), the true positive rate was 47 percent due to the imbalance between the class of cells targeted by poachers (69 out of 1,014) compared to the class of cells not targeted by poachers. To correct for this, we applied the synthetic minority oversampling technique (SMOTE) on the minority class,[6] obtaining a true positive rate of 67 percent, a true negative rate of 94 percent, and an overall 91 percent accuracy rate.

Rhino Behavior Model

In order to protect the rhinos, we must have a clear idea of where they are. We used the same features described earlier to predict rhino locations, and we learned interesting rules about rhino behavior.

For example, rhinos tend to visit cells *c* that have AVGELEV of less than 387 meters and STEEPNESS of less than 5 meters. (Support is 10.06 percent, confidence is 85 percent, negative confidence is 20.25 percent, and lift is 3.05.)

Simply put, from the 283 cells (of 1,014) that had documented rhino visits in our datasets, rhinos tended to visit cells based on the elevations and steepness of those cells. We derived several rules from the data that included minor variations of these preconditions but did not produce better support, confidence, negative confidence, or lift.

We also tried to predict the percentage of rhinos in a cell *c* during the next day (that is, one-day-ahead predictions). Using a tenfold cross validation using 80 percent of our data for training and 20 percent for validation, we developed regression models to predict these numbers. We tested linear regression, polynomial regression (with degrees 2, 3, and 4), support vector regression, and Gaussian process regression. We then compared our predictions with the actual numbers in the validation dataset (see Figure 1).

Gaussian process regression yielded the best results. When we compared the Pearson correlation coefficient between the predicted numbers and the actual numbers, we obtained a PCC of 0.79, showing a strong ability to

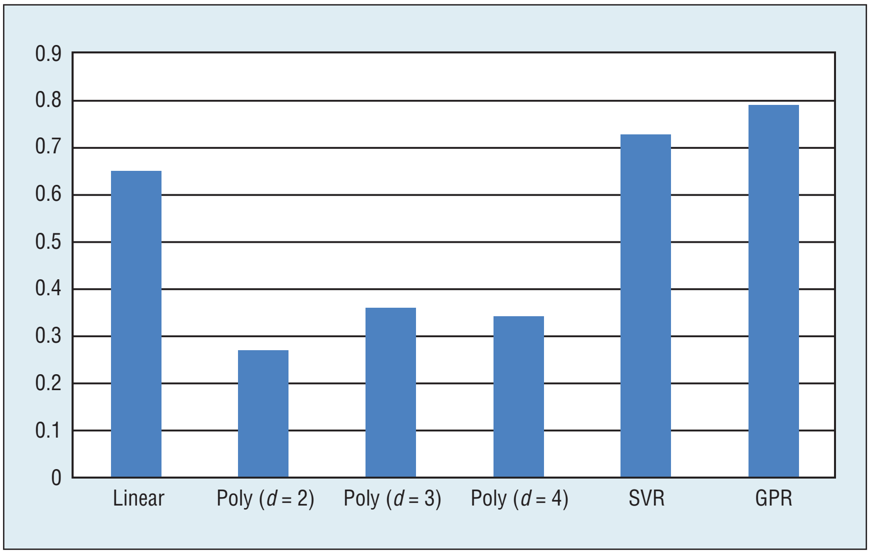


Figure 1. Pearson correlation coefficient (PCC) of the number of rhinos present in a cell (in reality) versus the number of rhinos predicted by our regression-based behavior model. Our model achieves high accuracy when used in conjunction with Gaussian process regression (GPR). (SVR: support vector regression.)

predict the number of rhinos visiting a given cell during a given day.

Artificial intelligence has produced world-class research in areas such as planning and scheduling, constraint solving, and other kinds of decision making. However, real-world behaviors can often derail the best performing (in the lab) planning and scheduling programs because the agents involved (humans, such as poachers, and animals, such as rhinos) could behave in unique and unanticipated ways. Developing optimal algorithms to make decisions in the presence of objective data and behavioral models is an area that needs considerable effort in AI over the next decade.

This article presents one such example in the context of a major international problem—trafficking in rhino horn. Our companion paper presents one decision-making problem,[3] which is coordinating drone flights and ranger patrols on the ground to maximize the expected number of rhinos that are protected. Using our South African data, we show that an algorithm that uses poacher behavior models (over and above a rhino behavior model) protects 20 percent more rhinos than using just the rhino behavior model. In South Africa alone, this suggests that predictive analytics can help protect far more animals than are currently being protected, with no significant increase in funding. We end on a note of caution—in other countries with varying species of rhinos, geographies, and sociocultural conditions (such as Kaziranga), the situation might differ. Nevertheless, this article shows that data-driven, behavioral-model-based predictive analytics could help save at least a few rhinos. ■

References

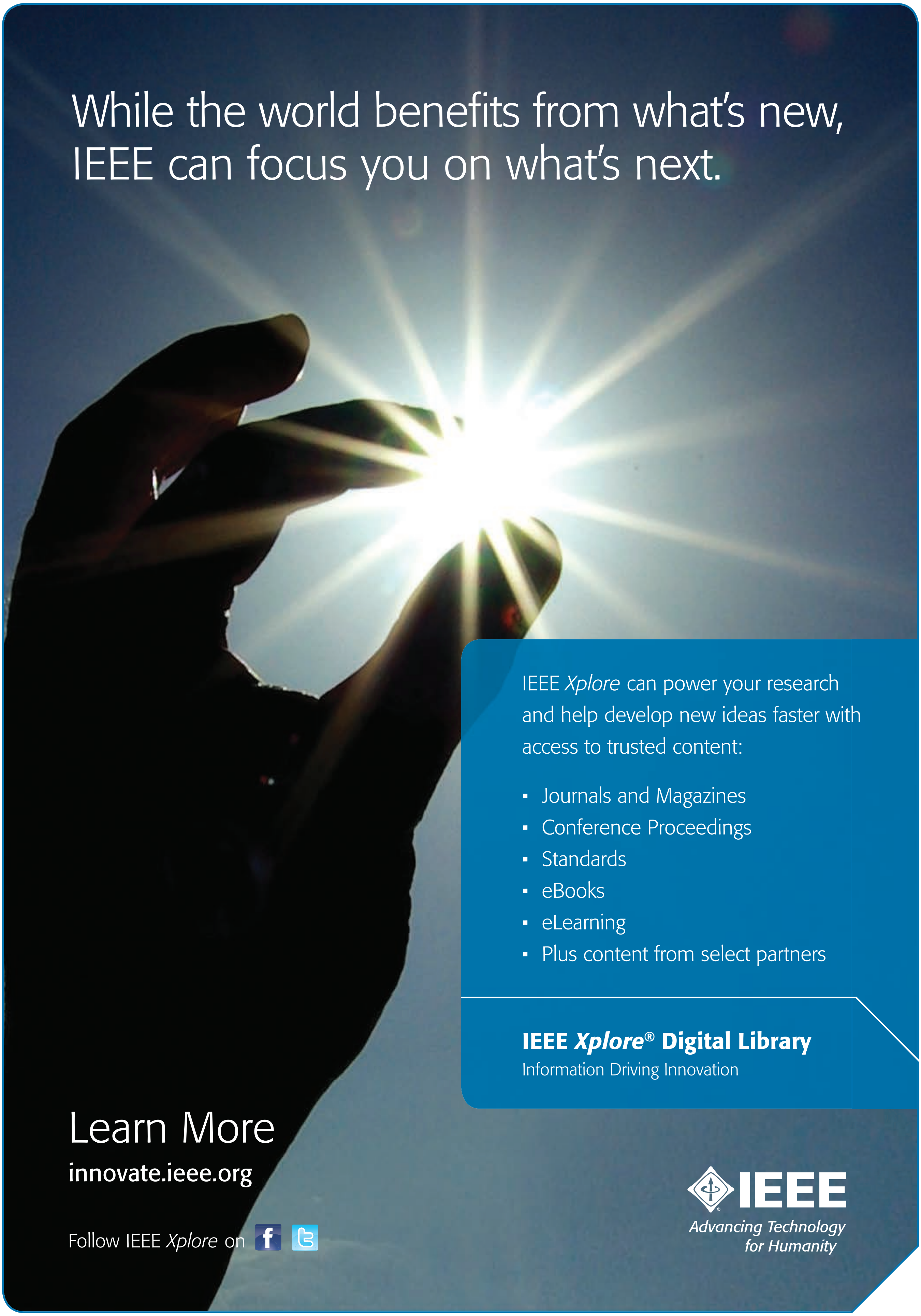
1. J.R. Platt, "1,215: The Record Number of Rhinos Poached in 2014," *Scientific Am.*, blog, Jan. 2015; http://blogs.scientificamerican.com/extinction-countdown/2015/01/23/record-number-rhinos-poached.
2. G. Guilford, "Why Does a Rhino Horn Cost $300,000? Because Vietnam Thinks It Cures Cancer and Hangovers," *The Atlantic*, 15 May 2013; www.theatlantic.com/business/archive/2013/05/why-does-a-rhino-horn-cost-300-000-because-vietnam-thinks-it-cures-cancer-and-hangovers/275881.
3. N. Park et al., "APE: A Data-Driven, Behavioral Model Based Anti-Poaching Engine," under review, *IEEE Trans. Computational Social Systems*.
4. R. Agrawal, T. Imieliński, and A. Swami, "Mining Association Rules Between Sets of Items in Large Databases," *ACM SIGMOD Record*, vol. 22, no. 2, 1993, pp. 207–216.
5. V.S. Subrahmanian and J. Ernst, *Method and System for Optimal Data Diagnosis*, US patent 7,474,987, to Univ. of Maryland, Patent and Trademark Office, 2009.
6. N.V. Chawla et al., "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artificial Intelligence Research*, vol. 16, no. 1, 2002, pp. 321–357.

Noseong Park is a PhD student in computer science at the University of Maryland. Contact him at npark@cs.umd.edu.

Edoardo Serra is a postdoctoral fellow at the University of Maryland Institute for Advanced Computer Studies (UMIACS). Contact him at eserra@umiacs.umd.edu.

V.S. Subrahmanian is a professor in the Department of Computer Science, director of the Lab for Computational Cultural Dynamics, and director of the Center for Digital International Government at the University of Maryland. Contact him at vs@cs.umd.edu.

This article originally appeared in IEEE Intelligent Systems, *vol. 30, no. 4, 2015.*


While the world benefits from what's new,
IEEE can focus you on what's next.
IEEE Xplore can power your research and help develop new ideas faster with access to trusted content:
▪ Journals and Magazines
▪ Conference Proceedings
▪ Standards
▪ eBooks
▪ eLearning
▪ Plus content from select partners
IEEE Xplore® Digital Library
Information Driving Innovation
Learn More
innovate.ieee.org
Follow IEEE Xplore on
IEEE
Advancing Technology for Humanity

Editor: Carole Goble • carole.goble@manchester.ac.uk

Smart Cities' Data: Challenges and Opportunities for Semantic Technologies

Mathieu d'Aquin • *The Open University, UK*

John Davies • *British Telecommunications, UK*

Enrico Motta • *The Open University, UK*

How can we innovate smart systems for smart cities, to make data available homogeneously, inexpensively, and flexibly while supporting an array of applications that have yet to exist or be specified?

Smart cities use technology to improve their services' efficiency and their inhabitants' lives.[1] Such use of technology is common in many other areas (including business, education, and government), but a fundamental difference with cities is the variety of both the types of technologies used, and the scenarios in which they can be applied, not to mention a highly diverse and distributed set of data sources.

The "smart city technology stack," if ever such a thing existed,[2] would typically include aspects of networking technologies, sensors and physical devices, communication technologies, cyber-physical systems, information systems, data management, federation and distribution, open data, analytics, visualization, machine learning, and many more components. Similarly, the notion of a smart city is one that encompasses a vast diversity of domains — from energy consumption to transport and mobility, and the management of city infrastructures or understanding citizens' opinions and behaviors. The consequence of all this is that a smart city is at best a system of systems,[3] and in most cases, many systems of systems. Here, we look at the role of semantic technologies in smart cities, using examples from two intiatives — MK:Smart and HyperCat.

Smart Cities' Linked Data and Semantic Web Technologies

The complexity and diversity of smart cities is one of the main reasons why they've emerged as key use cases for linked data and Semantic Web technologies recently (see the series of workshops "Semantics for Smart Cities," http://kat.ee.surrey.ac.uk/wssc/index.html; and "Semantic Cities," http://research.ihost.com/semanticcities14/). Linked data enable integrating data into a common, browsable, and accessible "conceptual" graph, while leaving data distributed and managed in different systems, under the control of different contributors. The use of linked data technologies has been shown effective in many cases where information from different sources must be put together in a generic way, to enable a variety of applications, without the need to encode the constraints of the applications in the data model. Semantic Web technologies add to this the ability to apply meaningful data models, in the form of shared Web vocabularies and ontologies, both to improve interoperability between systems (which might share the same meaning, but model it in a different way), and to enable a higher level of data analysis (see for example the work of Freddy Lécué and his colleagues[4]).

MK:Smart (http://mksmart.org) is a large collaborative initiative, partly funded by the Higher

 2376-113X/15/$31.00 © 2015 IEEE

Education Funding Council for England (HEFCE) and led by the Open University with British Telecommunications (BT) as a major industrial partner. It aims to develop innovative solutions to support economic growth in Milton Keynes (MK), a new town in Buckinghamshire, UK. Central to the project is the creation of a state-of-the-art "MK Data Hub" that supports acquiring and managing vast amounts of data relevant to city systems from a variety of data sources. These include data about energy and water consumption, transport data, data acquired through satellite technology, social and economic datasets, and crowdsourced data from social media or specialized apps. The idea is to build a common facility to efficiently manage, integrate, and re-deliver such data for applications and services to rely on, reducing development costs for all of these applications, and enabling intelligent data processing mechanisms (mining, analytics, aggregation, alignment, and linking) at the scale of the entire city, in a common data infrastructure.

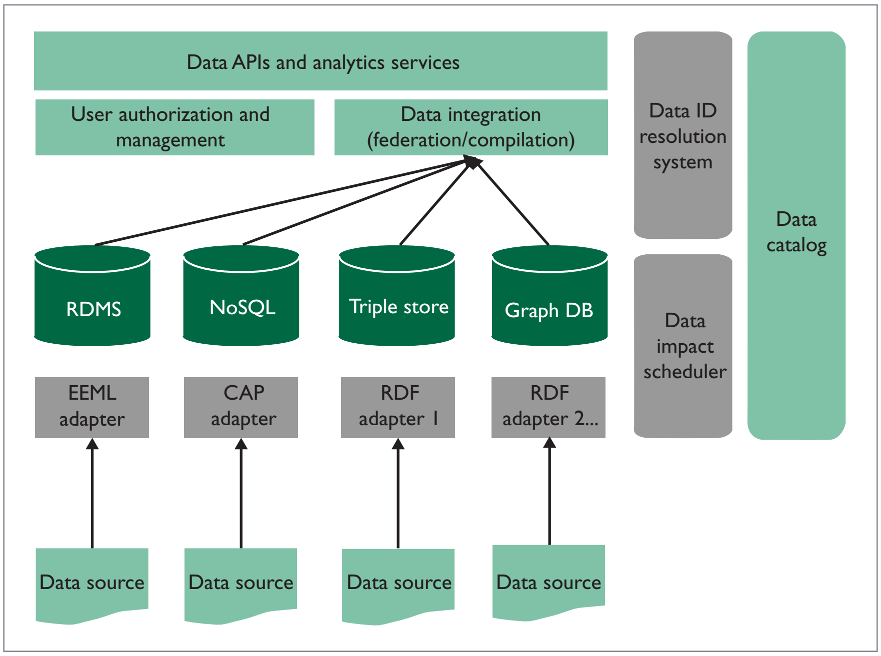


Figure 1. Architectural overview of the MK Data Hub developed in the MK:Smart project. One of the choices made in this architecture's design is to avoid relying on one single point of storage, to make maintenance and management of smaller storage components distributed among dedicated servers easier and more robust.

Although this might sound like a perfect application for a data warehouse,[5] there are a number of issues with traditional data integration approaches that make them harder to apply here. Indeed, the sheer number and variety of data sources means that a lot of effort must be put into transforming the data to fit not only the format of the data management infrastructure in place, but also to a common, global data model. In other words, the curation effort of such centralized methods becomes unsustainable. The other important aspect here is that, contrary to traditional approaches, the data management infrastructure of such a smart city data hub isn't built to support the requirements of a specific set of applications, services, or access patterns. The point of systems such as the MK Data Hub is to enable innovation, by making data available in a homogeneous, inexpensive, and flexible manner, supporting applications that might not be specified yet. But the data's value isn't entirely known at the time and by the people who are building the infrastructure to share, use, and consume it.

This is naturally where linked data and the Semantic Web can provide a technological advantage by supporting an approach to loosely integrate data that are kept as close as possible to their origin without having to encode models and schemas in the system, thereby constraining the data's applicability. To illustrate this point, in Figure 1 we summarize the approach taken in the MK Data Hub. As you can see, one of the choices made in the design of this architecture is to avoid relying on one single point of storage (or even one single storage technology). The reason for this choice is that maintenance and management of smaller storage components distributed among dedicated servers is easier and more robust than it would be with a unique warehouse when having to deal with thousands of datasets, with the sources of most of them being out of the data hub operators' control. Also, the development costs of import pipelines are reduced when we can choose the most appropriate storage format among a number of options for each data feed to be considered. Each data source might rely on a different format, a different mode of transfer, have different constraints attached to it, a different update rate, and so on, which we deal with through simple, lightweight, self-contained data-import pipelines.

We consider this approach a "lazy" method to data integration, because the data are only integrated at the last stage of going through the data hub, meaning that each data source/feed remains isolated in the infrastructure, and it's therefore easier to manage, up to the point of re-delivery. At this point, data are aggregated through a process akin to "dynamic compilation"[6] where, following the linked data principles, global identifiers for data entities are reconciled with their local identifiers in each of the datasets, and query federation is applied to obtain information aggregated and refactored into a common virtual data model at the time of data access.

Semantic/Linked Data Cataloging for a Smart City Data Hub

Similar to our aforementioned discussion, much of the effort in applying linked data and Semantic Web principles have so far focused on data interoperability at the data content level: figuring out how to make it easier to use different datasets together (from different places and using different data models). However, one of the consequences of taking such approaches and using linked data principles is that different data sources contribute to the information flowing through the hub, and information about these data sources must be tracked, too. Although this isn't a new problem, it especially becomes amplified in the context of smart cities, where many of such data hubs are likely to exist (at least one for each city), each covering datasets from thousands of different sources. In other words, in such contexts, making the information about these datasets manipulable, understandable, and exploitable becomes necessary, because traditional (mostly manual) approaches applied in common closed and restricted corporate information systems are insufficient to deal with the implied scale and complexity. Then the datasets and data sources in a smart city data hub become data management issues, in which linked data and Semantic Web technologies have a role to play.

HyperCat (see www.hypercat.io) is a clear example of an initiative that aims to address this particular issue. In 2014, innovateUK (the UK's innovation agency) funded a program called the Internet of Things Ecosystem Demonstrator. Eight industry-led projects were funded to deliver IoT clusters. Each cluster focused on a different domain, such as smart transport (the BT-led www.stride-project.com), airports, smart homes, schools, and so on. Clusters centered around a data hub to aggregate and expose data feeds from multiple sensor types. A major program objective was to address interoperability, specifically on how interoperability could be achieved between data hubs in different domains: this led to HyperCat, a standard for representing and exposing the Internet of Things data hub catalogs[7] over Web technologies, to improve data discoverability and interoperability. The idea is to enable distributed data repositories (data hubs) to be used jointly by applications, making it possible to query their catalogs in a uniform machine-readable format. This enables the creation of "knowledge graphs" of available datasets across multiple hubs that applications can exploit and query to identify and access the data they need, in whatever data hub they reside. HyperCat's specification[8] achieved this by employing the same principles on which linked data and the Semantic Web are built: data accessible through standard Web protocols and formats (HTTPS, JSON, and so on), identifying resources through URIs, and establishing common, shared semantics for datasets' descriptors.

From this perspective, HyperCat represents a pragmatic starting point to solve the issues of managing multiple data sources, aggregated into multiple data hubs, through linked data and Semantic Web approaches. It incorporates a lightweight, JSON-based approach constructed from a technology stack used by a large population of Web developers, and as such offers a low barrier to entry, thereby encouraging smart city ecosystems' growth.

Each HyperCat catalog lists and annotates any number of URIs (which typically identify data sources), each having a set of relation-value pairs (metadata) associated with it. In this way, HyperCat allows a server to provide a set of resources to a client, each with a set of semantic annotations. There are a small set of core mandatory metadata relations that a valid Hypercat catalog must include: beyond this, implementers are free to use any set of annotations to suit their needs. One important non-mandatory relation (rdf:type) in the specification allows a data feed to be associated with a Resource Description Framework Schema (RDFS) class, thereby linking to the Linked Open Data Cloud. With HyperCat, developers can write applications that will access data from many hubs, aiming to break down the walls between today's vertical silos. A Hypercat developer community is emerging, with open source tools becoming available (https://hypercatiot.github.io).

As Figure 2 shows, using HyperCat, compliant hubs and applications can communicate to establish what data a hub holds and how to access that data. HyperCat provides a standard means for resource discovery, which enables an interoperable ecosystem.

However, this notion of a "smart city data hub" also represents a clear and practical instance of a challenge that linked data and semantic technologies are only beginning to address, which must be integrated into standards such as HyperCat in order for smart cities to continue growing: ensuring data's traceability and rights to use the data. Indeed, recent work in the linked data world has focused on creating shared vocabularies and ontologies for describing the provenance of data (see the Provenance Interchange Ontology [PROV-O][9]), as well as applying models of digital rights on data resources (see the Open Rights Digital Language [ODRL],[10] and some applications related to linked data[11,12]).

These works make it possible to include information about the data's origin in the datasets' metadata, along with the processes they might have gone through, as well as the policies that apply to the data — expressed in terms of the actions that are permitted, prohibited, and required by the data owner through the data license. What the smart city data hub concept is starting to demonstrate (as exemplified in the MK:Smart project[13]) is that managing and exploiting such information requires much more machine intelligence than we might expect initially.

Indeed, each of the thousands of datasets included in each of the dozens of data hubs that an application might have to interact with can use different sets of policies and licenses. These datasets' policies and licenses might represent different (and sometimes incompatible) constraints, so that as they're processed through the data hubs it could affect each one differently; the same is true with the application itself. In other words, while HyperCat provides a starting point and a base to address the issue of curating catalogs of datasets and data sources, such catalogs need to become much more sophisticated Semantic Web systems that can intelligently handle and make sense of the complexity of policies attached to the datasets they cover, of the data flows they're implementing, and of the interaction between these data flows and data policies. This might sound like a minor detail in comparison to the challenge of collecting, curating, and analyzing data at the scale of a city or multiple cities. But the reality is that the ability to handle, propagate, and enforce data policies across smart city data hubs might turn out to become a key bottleneck to scaling-up smart cities, which might only be overcome through intelligent, semantic mechanisms to automate the manipulation and reasoning upon such aspects.

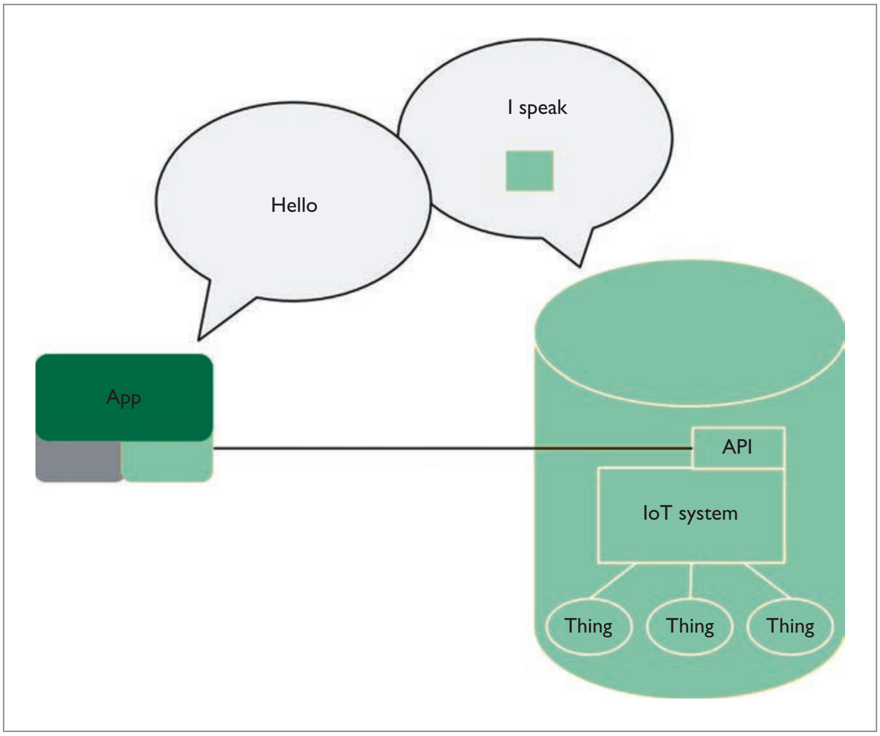


Figure 2. How applications interact with data hubs through HyperCat. Compliant hubs and applications can communicate to establish what data a hub holds and how to access that data.

The solution applied in MK:Smart is to develop this type of sophisticated catalog,[13] where data policies and licenses — as well as the data flows that relate to them[14] — are represented as machine-readable information. This, in turn, enables the implementation of complex inference rules[15] to support the automatic manipulation of such information in tasks, such as data discovery and policy validation.

As the MK Data Hub evolves, it becomes an example of how smart city data hubs are no longer only about enabling smart systems that intelligently exploit data, but also about being smart systems themselves that effectively manage a large number of diverse datasets (using linked data and Semantic Web principles), thereby providing intelligent support to consume such data.

References

1. A.M. Townsend, *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*, WW Norton & Company, 2013.
2. T. Nam, and T.A. Pardo, "Conceptualizing Smart City with Dimensions of Technology, People, and Institutions," *Proc. 12th Ann. Int'l Digital Government Research Conf.: Digital Government Innovation in Challenging Times*, 2011, pp. 282–291.
3. J. Boardman and B. Sauser, "System of Systems — The Meaning of," *Proc. IEEE/SMC Int'l Conf. System of Systems Eng.*, 2006; doi:10.1109/SYSOSE.2006.1652284.
4. F. Lécué et al., "Predicting Severity of Road Traffic Congestion Using Semantic Web Technologies," *The Semantic Web: Trends and Challenges*, Springer, 2014, pp. 611–627.
5. B. Devlin and L.D. Cote, *Data Warehouse: From Architecture to Implementation*, Addison-Wesley Longman Publishing, 1996.

6. A. Adamou and M. d'Aquin, "On Requirements for Federated Data Integration as a Compilation Process," *Proc. 2nd Int'l Workshop on Dataset Profiling and Federated Search for Linked Data at 12th European Semantic Web Conf.*, 2015; http://ceur-ws.org/Vol-1362/PROFILES2015_paper4.pdf.
7. J. Davies and M. Fisher, "Internet of Things: Why Now?" *J. Inst. of Telecomm. Professionals*, vol. 9, no. 3, 2015, pp. 36–42.
8. P. Beart, T. Jaffrey, and J. Davies, *HyperCat 2.0 Specification*, HyperCat Foundation, 2015; http://wiki.1248.io/doku.php.
9. T. Lebo et al., *PROV-O: The PROV Ontology*. W3C recommendation, 2013.
10. R. Iannella, *Open Digital Rights Language (ODRL) Version 1.1*, W3C note, 2012.
11. S. Villata and F. Gandon, "Licenses Compatibility and Composition in the Web of Data," *Proc. 3rd Int'l Workshop on Consuming Linked Data.*, 2012; http://ceur-ws.org/Vol-905/VillataAndGandon_COLD2012.pdf.
12. E. Daga et al., "A Bottom-Up Approach for Licences Classification and Selection," *Proc. Int'l Workshop on Legal Domain and Semantic Web Applications at 12th Extended Semantic Web Conf.*, 2015; http://cs.unibo.it/ledaswan2015/papers/daga-et-al-ledaswan2015.pdf.
13. M. d'Aquin et al., "Dealing with Diversity in a Smart-City Datahub," *Proc. 5th Workshop on Semantics for Smarter Cities*, 2014; http://ceur-ws.org/Vol-1280/paper8.pdf.
14. E. Daga et al., *Describing Semantic Web Applications through Relations between Data Nodes*, tech. report kmi-14-05, Knowledge Media Inst., Open Univ., 2014; http://kmi.open.ac.uk/publications/techreport/kmi-14-05.
15. E. Daga et al., "Propagation of Policies in Rich Data Flows," *Proc. 8th Int'l Conf. Knowledge Capture*, 2015, pp. 5:1–5:8; doi:10.1145/2815833.2815839.

Mathieu d'Aquin is a senior research fellow at the Knowledge Media Institute of the Open University, UK. His research focuses on semantic technologies, the Semantic Web, and linked data. D'Aquin has a PhD in artificial intelligence from the University of Nancy. Contact him at mathieu.daquin@open.ac.uk.

John Davies is the chief researcher in British Telecommunications' Future Business Technology Research Practice. His interests include the application of semantic and Internet of Things technologies to smart cities, smart transport, business intelligence, and information integration. Davies has a PhD in artificial intelligence from the University of Essex. Contact him at john.nj.davies@bt.com.

Enrico Motta is a professor in knowledge technologies at the Open University. His research focuses on large-scale data integration, pattern extraction, and visualization to support users in making sense of large amounts of data. Motta has a PhD in artificial intelligence from the Open University. Contact him at enrico.motta@open.ac.uk.

This article originally appeared in IEEE Internet Computing*, vol. 19, no. 6, 2015.*

stay connected.

Keep up with the latest IEEE Computer Society publications and activities wherever you are.

IEEE computer society

| @ComputerSociety
| @ComputingNow

facebook | facebook.com/IEEEComputerSociety
| facebook.com/ComputingNow

LinkedIn | IEEE Computer Society
| Computing Now

YouTube | youtube.com/ieeecomputersociety

IEEE

IEEE computer society

intel
IEEE CS Software Professional Developer
Platinum
2016

Get the Recognition You Deserve

The 2016 Platinum IEEE Computer Society/Intel Software Developer Award recognizes the most talented software developers in the world.

3 WINNERS will be selected. The 3 developers with the highest exam scores will each receive the Award along with $3000. Winners are also invited to the IEEE CS Annual Awards Banquet in Atlanta, GA, where they will be honored by computing industry leaders.

Complete the Software Professional Developer Certification prior to May 1, 2016, and your test scores will automatically be entered to qualify you for these distinguished awards.

Validate your expertise. Gain global recognition. Advance your career.

FOR COMPLETE INFORMATION ON HOW TO ENTER, GO TO
www.computer.org/intel-award

Editor in Chief: **Diomidis Spinellis**
Athens University of Economics
and Business, dds@computer.org

Extending Our Field's Reach

Diomidis Spinellis

I RECENTLY COLLABORATED with a digital-typography expert to create a formatting style for a publishing project. I thought we had agreed to work together through GitHub, which would let us share the current version of the manuscript and easily integrate the LaTeX style files by merging their corresponding development branch. Instead, a few weeks later, I received a .zip archive file over email.

> Almost all software engineering processes can benefit industries that work with executable knowledge.

The style was beautiful, with great attention to detail and typographical niceties. The archive's contents were also interesting. It contained two directories with similar but not identical contents: one for all projects requiring the specific style and one slightly tailored to my needs. Each directory contained tens of third-party files (some in binary format), log files, documentation, and automatically generated files. The style's source code files also contained things that might trouble a software engineer: outdated or inconsistently indented comments, copy-and-pasted and commented-out code, and a few overly long lines. To be fair, some of the style code seemed to have been written more than two decades ago, and we all know how software ages.

Still, whenever I collaborate with publishers, I'm always saddened to see the opportunities for process improvement they miss by not using the best practices and tools we've established in software engineering. For example, instead of collaboration using simple text markup over an online revision control and review system, documents in diverse incompatible binary formats are shuttled back and forth over email and, yes, FTP, with changes and comments embedded obscurely (sometimes with typographical marks scribbled on the margins of scanned paper). This creates an integration nightmare, which is only partially controlled by draconian, inflexible change-management policies and heroic efforts of all the involved parties. Once a document leaves a specific stage—say, drafting, copyediting, or composition—there's no going back, and nobody can trace changes on an end-to-end basis. I know that some publishers use version control systems, but even there, due to the lack of build-process automation, such use often degenerates into that of a shared disk drive.

Picking on publishing would be wrong; other industries are also producing what's in effect software (ex-

IEEE Software Mission Statement
To be the best source of reliable, useful, peer-reviewed information for leading software practitioners—the developers and managers who want to keep up with rapid technology change.

 2376-113X/15/$31.00 © 2015 IEEE

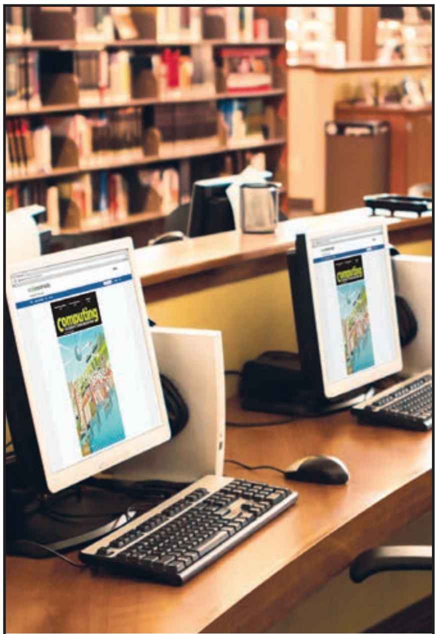
ecutable knowledge) but not treating it as such. As examples, consider 3D printing, numerical control of machine tools, filmmaking, pharmaceutical laboratory automation, and workflow management. Other activities, with more abstract output, include project planning and drafting laws and regulations. The output of these activities shares considerable similarities with software: laws are constantly revised and refer to parts of each other, similarly to subroutines. These examples don't include devices in which software (the kind that runs on a CPU) is taking an ever-increasing role, such as cars, planes, phones, medical devices, and TVs. There, the problems and missed opportunities are much less severe; trained software engineers typically perform and manage the development.

Although many industries have developed their own highly effective processes over the years, software engineering maintains an essential advantage. It has developed methods and tools that let even small teams manage extremely high complexity. For example, compare the nine million LOC in the Linux kernel (which often forms only a small part of a much larger software stack) with the few tens of thousands of components in a modern car. This advantage is important because the complexity in nonsoftware activities is also increasing inexorably. Films used to have a few hundred scenes and takes, which filmmakers could easily manage by writing on a clapperboard. In contrast, modern blockbusters might depend on tens of thousands of digital artifacts. Also, processes that skilled humans executed only a few times in the past (for example, creating a car model for aerodynamic testing) can now be easily rerun (for example, by a 3D printer) with the touch of a button.

Almost all software engineering processes can benefit industries that work with executable knowledge. Requirements engineering can improve analysis, specification, validation, and traceability (why do we drill this hole at the bottom of the engine block?). Software design can be essential to control complexity through modeling, abstraction, control of coupling and cohesion, decomposition, encapsulation, and separation of concerns.

Returning to the example I presented at the beginning, the two directories I received from the publisher should have been divided into a meaningful hierarchy (decomposition) and their contents united through some application of polymorphism. Construction techniques can be used for the promotion of agility, build automation, continuous integration, reuse management, and verification. Imagine a film director being always able to use continuous integration to view the most current version of a film as it develops over the months. Software-testing techniques can reduce waste and increase efficiency, while software-inspired maintenance activities can increase a product's or process's longevity.

I believe that configuration management tools and techniques are the most productive way through which software engineering can affect other fields. The powerful tools and platforms we've developed (think of Git and GitHub) allow the effortless sharing of work over distance and time zones, the documentation and traceability of changes, the integration of issue and change management, commenting on each other's work, the organized development of separate product lines, and the



It's already at your fingertips

Computing in Science & Engineering (CiSE) appears in the IEEE Xplore and AIP library packages, so your institution is bound to have it.

pain-free merging of work items and branches. Engineers in other fields often look at these feats as alien technology. Some people, though, are recognizing the potential; consider the appearance of French civil code and German laws on GitHub (https://github.com/steeve/france.code-civil and https://github.com/bundestag/gesetze).

Exporting our hard-earned knowledge to other fields won't be easy. Each has distinct goals, competencies, values, and traditions. To communicate effectively, we must develop a shared vocabulary and way of thinking. Perhaps education is the easiest path: train practitioners from other disciplines to think and act as software engineers.

Software engineering has benefited mightily from research in fields ranging from electrical engineering and physics to mathematics and management science. It has changed our world beyond recognition by putting the artifacts it produces on billions of devices. Now, the time has come to transform our world in another way, by giving back to science and technology the knowledge software engineering has produced in the past half century.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.

This article originally appeared in IEEE Software, *vol. 32, no. 6, 2015.*

Call for Articles

IEEE Software seeks practical, readable articles that will appeal to experts and nonexperts alike. The magazine aims to deliver reliable, useful, leading-edge information to software developers, engineers, and managers to help them stay on top of rapid technology change. Topics include requirements, design, construction, tools, project management, process improvement, maintenance, testing, education and training, quality, standards, and more.

Author guidelines: www.computer.org/software/author.htm
Further details: software@computer.org
www.computer.org/software

IEEE Software

CONFERENCES
in the Palm of Your Hand

IEEE Computer Society's Conference Publishing Services (**CPS**) is now offering conference program mobile apps! Let your attendees have their conference schedule, conference information, and paper listings in the palm of their hands.

The conference program mobile app works for **Android** devices, **iPhone**, **iPad**, and the **Kindle Fire**.

For more information please contact **cps@computer.org**

IEEE
IEEE computer society
CPS Conference Publishing Services

Multimedia Big Data Computing

Wenwu Zhu, Peng Cui, and Zhi Wang
Tsinghua University, China

Gang Hua
Steven Institute of Technology

With the proliferation of the Internet and user-generated content, and the growing prevalence of cameras, mobile phones, and social media, huge amounts of multimedia data are being produced, forming a unique kind of big data. Multimedia big data brings tremendous opportunities for multimedia applications and services—such as multimedia searches, recommendations, advertisements, healthcare services, and smart cities. The need to compute such massive datasets is transforming how we deal with multimedia computing.

Researchers have studied some of the problems in big data computing (see the related sidebar), but multimedia big data has its own characteristics related to multimodality, real-time information, quality of experience, and so on. For example, some multimedia learning applications, games, or 3D rendering might require GPU processing. Consequently, methods for general big data might not directly apply to multimedia big data.

Compared to approaches of general text-based big data computing, multimedia big data computing faces additional compression, storage, transmission, and analysis challenges in terms of

- organizing unstructured and heterogeneous data,
- dealing with cognition and understanding complexity,
- addressing real-time and quality-of-service requirements, and
- ensuring scalability and computing efficiency.

Here, we consider these technical challenges and the related scientific problems for multimedia big data computing, introducing various research directions and emerging technologies.

The Multimedia Life Cycle

The emergence of big data computing is having a profound effect on the entire life cycle of multimedia content. Figure 1a shows the typical multimedia life cycle, which comprises acquisition, storage, processing, dissemination, and presentation.

In recent decades, the availability of low-cost commodity digital cameras and camcorders has sparked an explosion of user-generated media content. Most recently, cyber-physical systems have started offering a new type of data acquisition through sensor networks, significantly increasing the volume and diversity of media data.[1] Riding the Web 2.0 wave and social networks, digital media content can now be easily shared through the Internet, including via social networks. The huge success of YouTube demonstrates the popularity of "Internet" multimedia; similarly, social multimedia has had great success thanks to social networks such as Facebook and Twitter.

In the early stage, media storage, processing, and dissemination were relatively small in scale—usually at the level of kilobytes. Now, the data scale is often at the terabyte or even petabyte level. The collected datasets are so large and complex that it becomes difficult to process using traditional media data processing technology.

However, multimedia big data provides great opportunities. Both the scale and richness of the data—in terms of content, context, users and crowds, and so on—provide more opportunities to build better computational models to mine, learn, and analyze enormous amounts of data. Moreover, multimedia big data algorithms require "massively parallel software running on thousands of servers distributively."[2]

A typical multimedia big data computing life cycle consists of moving from data to information, from information to knowledge, from knowledge to intelligence, and from intelligence to decision, as depicted in Figure 1b. First, we need to process the collected multimedia raw data into information, creating multimedia knowledge and insight. When we combine this output with human or user knowledge, it can

 Published by the IEEE Computer Society 2376-113X/15/$31.00 © 2015 IEEE

Related Work in Big Data Computing

Various researchers have studied the challenges of big data computing. Xindong Wu and his colleagues presented a Heterogeneous, Autonomous, Complex, Evolving (HACE) theorem that characterizes the features of the big data revolution and proposes a big data processing model from the data mining perspective.[1] Philip Russom and his colleagues explained the concept, characteristics, and needs of big data and different offerings available in the market to explore unstructured large data.[2] Han Hu and his colleagues presented a literature survey and system tutorial for big data analytics platforms, aiming to provide an overall picture for nonexpert readers and instill a do-it-yourself spirit for advanced audiences to customize their own big data solutions.[3] Puneet Singh Duggal and Sanchita Paul suggested various methods for catering to the problems at hand through a Map Reduce framework over the Hadoop Distributed File System.[4] Stephen Kaisler and his colleagues have also analyzed the issues and challenges for big data analysis and design.[5] Changqing Ji and his colleagues introduced several big data processing technics from both system and application aspects.[6]

References

1. X. Wu et al., "Data Mining with Big Data," *IEEE Trans. Knowledge and Data Eng.*, vol. 26, no. 1, 2014, pp. 97–107.
2. P. Russom et al., "Big Data Analytics," *TDWI Best Practices Report, Fourth Quarter*, 2011.
3. H. Hu et al., "Toward Scalable Systems for Big Data Analytics: A Technology Tutorial," *IEEE Access*, vol. 2, 2014, pp. 652–687.
4. P.S. Duggal and S. Paul, "Big Data Analysis: Challenges and Solutions," *Proc. Int'l Conf. Cloud, Big Data and Trust 2013*, 2013.
5. S. Kaisler et al., "Big Data: Issues and Challenges Moving Forward," *Proc. 46th Hawaii Int'l Conf. System Sciences* (HICSS), 2013, pp. 995–1004.
6. C. Ji et al., "Big Data Processing in Cloud Computing Environments," *Proc. Int'l Symp. Parallel Architectures, Algorithms and Networks* (I-SPAN), 2012, pp. 17–23.

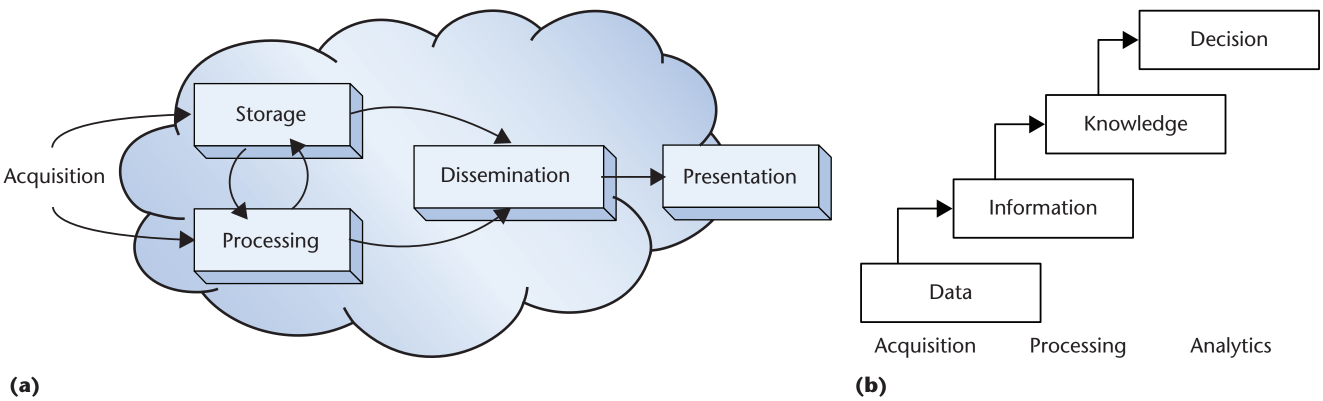


Figure 1. The emergence of big data computing is affecting the life cycle of multimedia content: (a) the typical multimedia life cycle and (b) the typical multimedia big data computing life cycle.

be used to make decisions. However, with this big data scale comes tremendous challenges.

Challenges

Compared to approaches of general text-based big data computing, multimedia big data computing faces the following fundamental technical challenges related to processing, storage, transmission, and analysis.

Unstructured and heterogeneous data. Multimedia big data is unstructured, heterogeneous, and multimodal, which makes multimedia big data representation and modeling difficult. For example, how do we turn unstructured multimedia data into structured data? How can we represent or model multimedia big data coming from different sources or spaces (cyber, physical, and social)?

Cognition and understanding complexity. Computers can't easily understand multimedia big data, mainly due to the semantic gap between low-level features and high-level semantics. Moreover, some multimedia big data is evolving with time and space.

Real-time and Quality of Experience (QoE) requirements. Multimedia big data applications and services are typically real time, so to address QoE requirements, we need real-time streamed/online, parallel/distributed processing for analysis, mining, and learning.

Scalability and efficiency. Multimedia big data systems need large-scale computation, so they must optimize computation, storage, and networking/communication resources. Such systems also need online/streamed and parallel/distributed algorithms. In addition, GPU computing for multimedia big data computing brings further challenges.

Scientific Problems

The four fundamental challenges just discussed lead to four corresponding scientific problems.

Representation and modeling. How can we establish the representation and modeling for unstructured, heterogeneous, and multimodal multimedia big data?

Deep and crowd computing. How do we perform data-driven deep computing (including mining and learning) to effectively analyze data? How can we exploit crowdsourcing jointly with data-driven analysis for multimedia cognition?

Streamed or online computing. How do we perform streamed/online processing for the entire multimedia big data in a parallel/distributed way, so as to make multimedia processing, analysis, mining, and learning real-time while satisfying QoE requirements?

Computing, storage, and communication optimization. How do we design a new multimedia computing architecture to optimize computation, storage, and network/communication for multimedia big data computing? How can we efficiently use GPU-powered servers for multimedia big data computing?

Addressing the Issues

Addressing these fundamental challenges and scientific problems for multimedia big data computing will require implementing effective approaches throughout the multimedia life cycle. Next, we look at each stage of the multimedia big data computing life cycle to identify ways of addressing these issues.

Data Acquisition

In addition to acquiring multimedia data from the Internet and Internet of Things (IoT) (for example, in the form of user-generated content and camera data), the emergence of online social networks makes it possible to collect multimedia data from individuals acting as sensors of the real world.

Raian Ali and his colleagues have proposed *social sensing,* in which users act as monitors and provide information needed at runtime.[3] Even before the conceptualization of social sensing, Anmol Madan and his colleagues had already tried using collected information from users to detect epidemiological behavior change—for example, to predict a person's health status using data collected from his or her cellphone.[4] On the other hand, the rapid development of wireless networks and mobile devices makes most collected multimedia data personal. Videos, images, and other kinds of multimedia data are increasingly correlated with individuals rather than with public groups.

Today, personal data can be collected by wearable sensors that can feed 24/7 data streams, acting as an important type of multimedia data source. This huge amount of multimedia data raises the following questions regarding data reduction (or compression) and data representation.

Data Reduction/Compression

The volume of multimedia big data must be reduced for efficient storage and communication. Multimedia data reduction refers to sampling the (massive) dataset so that it can be computed with limited computing resources. Multimedia data compression refers to reducing the raw data size for storage or communication.

Feature-transformation-based data reduction. The goal is to reduce numerical data using common signal processing or transform techniques. Compressive sensing[5] and Wavelet Transform are two approaches for big data reduction.[6]

Analysis-aware compression. Multimedia coding can be performed prior to analysis such that the compression-aware analysis can exploit the coding. For example, you might achieve compression using a feature descriptor.[7] During compression, you compress feature descriptions along with the images or videos. The compression, storage, and transmission of local feature descriptors of image and video applications can then later be used for computing, such as for a content-based search (or visual search).[7]

Cloud-based compression. Data compression conducted on the (sensor) client side can effectively save storage space by compressing the data after data generation but before storage. On the cloud side, cloud-based big data compression can take advantage of data correlation and background similarity.

Data Representation

Multimedia data representation refers to a mathematical structure in which you can model the data for later analysis. Because multimedia data comes from multiple sources, it tends to have disparate representations for each source, or sometimes a common representation is needed for multimodal analysis. These interpretations are usually described by both structural and descriptive metadata. Example data representation is referred to as feature-based data representation (such as scale-invariant feature transform).

Multimedia big data representation consists of the following approaches.

Feature-based data representation. Because multimedia big data is often multimodal, sometimes we can find a common feature space to represent the data—namely, feature-based representation. Feature-selection-based data representation aims to find the best representational data among all possible feature combinations.

Learning-based representation. Today's data comes from heterogeneous sources, such as cyber-physical-social spaces, so a common explicit feature space cannot be easily found. A new approach is to find implicit "hidden space" data representation for multimodal and heterogeneous data. Such representation is also called a learning-based (rather than hand-crafted) representation.

Many machine-learning methods have been proposed to represent multimedia data. For example, deep learning represents one breakthrough in representation learning. There are a series of deep learning methods to perform multimedia data representation, such as Deep Boltzmann Machines[8] and Deep Autoencoder.[9] Such methods aim to learn high-level representations from low-level features using a set of nonlinear transformations and have achieved the state-of-the-art for various tasks.

Computation-oriented representation. A key factor in multimedia data representation is to reduce computation, which requires understanding the relationship between the representation and its computational complexity. For example, in a content-based image search, hashing-based indexing methods map images represented by high-dimensional raw features into a Hamming space, where the images are represented by short binary hashing codes. By retrieving samples whose hash codes are within a small Hamming distance of the hash codes of queries, these methods can implement an efficient search. In addition, the compact hash codes can dramatically reduce the required storage space.

Data Processing and Analysis

After acquisition and storage, the next step in the multimedia life cycle is multimedia big data processing and analysis.

Multimodal Analysis

Multimedia data analysis has largely focused on how to fuse the information from the different media modalities together to form a coherent decision. However, issues arise when data entries lack data modality.

For example, many images shared on the Internet now come with geotags. Furthermore, beyond traditional RGB information, many images have additional information such as depth (taken, for example, from advanced sensors such as Kinect). However, there is a large backlog of images that do not possess additional sensing information, which present a problem for algorithms designed to operate using both the RGB and the geotags and/or depth information.

This calls for media analysis technologies that are robust despite missing data modalities, leading to the notion of cross-modality media analysis. From a machine-learning viewpoint, the problem can be formulated as follows: in the training phase, we have full access to data samples with different modalities, but in the testing phase, we might encounter data samples that are either systematically or randomly missing certain views of information. Qilin Zhang and his colleagues recently studied this problem and proposed a common latent space approach.[10] As shown in Figure 2, the basic idea of this type of approach is that in the training phase, a common latent space is identified for all data modalities, and all data samples from the different modalities are projected into the same space for reasoning. Then, in the testing phase, even if some data modalities are missing, the remaining ones can still be projected into the common space for reasoning.

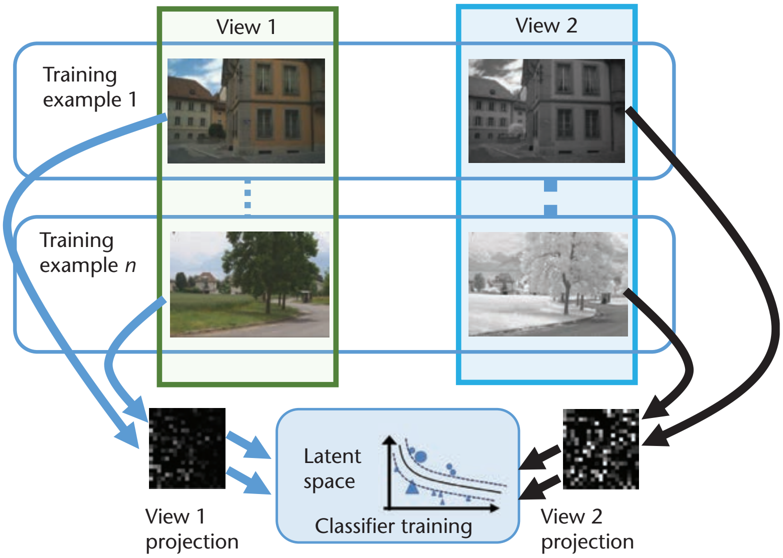


Figure 2. The common space model for cross-modal media analysis.

How to identify such common space is an open question. It could be conducted, for example, via canonical correlational analysis and its variants, or using more complicated methods, such as a deep neural network. When it comes to the problem of making decisions using multimedia big data, a fundamental issue is fusing the data from the different modalities. Previous methods largely focused on either pre-fusion at the feature level or late-fusion at the decision level.

Both fusion methods have their pros and cons. Feature-level fusion lets the decision algorithm potentially benefit from the correlational information across the two different feature modalities. However, finding a way to appropriately normalize the features from the different modalities is an open issue. Decision-level fusion often learns a weighted linear combination of the decision scores from each feature modality. It does not need to deal with the feature normalization issue. However, decision-level fusion might not be able to effectively leverage the correlational information across the different modalities.

We advocate a mid-level fusion scheme at the representation level, where we learn a compact intermediate-level representation to effectively capture the correlational information across the different modalities. Intuitively, such intermediate-level representation can effectively capture the correlational information across the different media modalities while suppressing the heterogeneity. For example, Zhenxing Niu and his colleagues proposed a visual topic network, which effectively learns an intermediate-level image representation from both visual features and sparse text tags, as illustrated in Figure 3.[11] This approach has shown consistent performance gain in terms of recognition accuracy, when compared with pre-fusion and late-fusion.

User-Centric Analysis

Although the rapid advances of multimedia computing technology have greatly facilitated users' information needs regarding multimedia data, it is still difficult for all multimedia applications (such as multimedia search and recommendation applications) to provide satisfactory results for users with different intentions. This is because there's a lack of understanding of users, which is more serious in multimedia applications than in nonmultimedia applications for two reasons.

First, multimedia applications often become more exploratory, and users are often interested in images or videos with a particular style, which is difficult to express and represent. Second, multimedia data is often used for entertainment when exploring a visual space, with no clear end goal.[1] How to discover users' latent intent from limited observed data is of paramount importance in improving multimedia search and recommendation performance. It resonates well with the idea that underpins *user-centric multimedia analysis,* where the user profiles, behaviors, and social networks are sensed, harnessed, and shared to adapt the results of general multimedia search and recommendation engines to be more consistent with user intent.

User intention modeling. User modeling is crucial to addressing the intention gap problem in multimedia search and recommendation. Feng Qiu and Junghoo Cho represented user interest by topics, and they proposed a method to learn user preferences from past query-click history in Google.[12] Eugene Agichtein and his colleagues proposed a method to learn the user interaction model to better predict the user's preference in terms of search results.[13] Jaime Teevan and her colleagues explored rich models for interest modeling by combining multiple resources, such as search-related information, user-relevant documents, and emails.[14] More recently, researchers have investigated user-

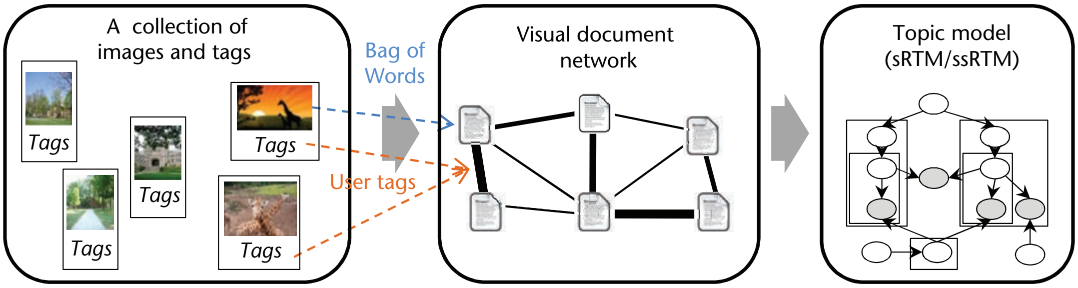


Figure 3. The visual topic network for representation-level fusion, where the sparse text tags link the different visual documents and contribute to the learning of the final representation of the images. (Source: Zhenxing Niu and his colleagues; used with permission.[11])

information interaction behavior patterns in social network environments.[15,16]

The interest-modeling problem is more challenging in the image domain due to the high-dimensional space and the semantic-gap problem. Marek Lipczak, Michele Trevisiol, and Alejandro Jaimes analyzed users' favorite behavior patterns in Flickr.[17] Xing Xie and his colleagues proposed detecting user interests from user-image interaction behaviors recorded by image browsing logs.[18] Yun Yang and her colleagues investigated the emotion prediction problem for individual users when watching social images.[19] Tags of images are mined to construct the topics and ontology to represent user preferences.[15] Similar to the problem that user intentions cannot be well represented by query words in an image search,[20] user interests in images cannot be well represented by tags. Visual factors, such as visual style and quality, eventually play important roles in user interest formation.

Social-sensed multimedia search. Personalized search has been studied for many years in the text domain. The main target has been to construct accurate and complete user profiles for re-ranking the search results by measuring the distance between the search results and user profiles.[21] More specifically, the user profiles were represented by an ontology[22] and topics,[12] which are mined from the metadata, search logs, and social media. Recently, some of these techniques have been transferred into a personalized image or video search, especially for image searches in Flickr.

Considering the special characteristics of social images, Dongyuan Lu and Qiudan Li proposed a co-clustering method to discover the latent interests of users and map the Flickr search results into the latent space to measure their matching degree. Kristina Lerman, Anon Plangprasopchok, and Chio Wong exploited user-generated metadata in the form of contacts and image annotations in Flickr to describe user interest, using them to re-rank the image search results in Flickr.[23]

Merging search engine and social media has clearly become a common trend in industry. For example, Google acquired YouTube and launched Google Plus; Yahoo acquired Flickr; and Facebook put forth efforts to develop search services with a Facebook-external scope. Much could be leveraged by integrating social media platforms with multimedia search systems, as shown in Figure 4. How to discover and represent user search intention from social media and seamlessly bridge these user intentions with multimedia search systems is a research issue in need of serious attention.

Social multimedia recommendation. Content-based filtering and collaborative filtering (CF) have been widely used to help users discover the most valuable information to them. Content-based filtering introduces the basic idea of studying an item's content to address the ranking problem. With the emergence of topic modeling techniques such as Latent Dirichlet allocation (LDA), recent content-based approaches[24] rank candidate items by how well they match the topical interest of the user. These methods represent users and items in fine granularity.

CF methods, consisting of memory- and model-based methods, are widely used. The memory-based approaches calculate the similarity between all users based on their ratings of items.[25] They represent users (or items) by the item sets (or user sets), which are often unstable and can only obtain good performance for

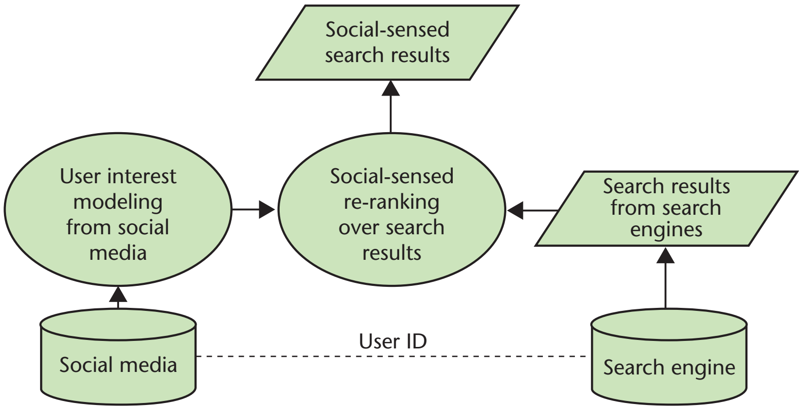


Figure 4. An illustration of the social-sensed image search.

active users or popular items. The model-based methods learn a model based on patterns recognized in the user ratings.

Several matrix factorization methods have recently been proposed.[26] The matrix approximation models all focus on representing the user-item rating matrix with low-dimensional latent vectors. Recognizing that influence is a subtle force that governs the dynamics of social networks, influence-based recommendation[27] involves interpersonal influence in social recommendation cases. Trust-based approaches[28] exploit the trust network among users and make recommendations based on the ratings of users who are directly or indirectly trusted.

Jiang and his colleagues proposed a probabilistic factor analysis framework, which fuses users' preference and social influence together.[14] Furthermore, they have also investigated the social recommendation problem in a multiple domain setting. Most of these works are based on traditional content-based filtering or CF-based methods, and their common goal is to embed social information into traditional methods to improve the recommendation accuracy. However, few authors have targeted the problem of how to learn a new common representation for users and items in social networks, which is indeed feasible and important for boosting social recommendation performance.

Human-in-the-Loop Analysis

Multimedia big data analysis is difficult, and many algorithms and systems, if running in a purely automated fashion, would not be able to achieve the level of performance required for practical use. This has motivated researchers to explore a hybrid human-computer method for content analysis tasks. The Visipedia project[29] is one pioneering project in this area, where interactive human inputs and advanced computational algorithms are tightly integrated to solve the content analysis problem.

This type of hybrid human-computer system is further catalyzed by crowdsourcing, where cheap online human labors can be exploited with a small fee on a per-input basis. Figure 5 provides an illustration of such hybrid human-computer systems. There are several issues that need to be carefully modeled when considering crowdsourcing-based human inputs. First of all, human input from crowdsourcing could be very noisy, so it should be carefully modeled. Often, inputs from several online workers are solicited to enable consensus-based analysis or to model the quality of the workers.

The second question is how and when human input should be engaged. It cannot be too frequent, because the cost would be high and the response time long. In this regard, uncertainty- and confidence-based reasoning would be critical, because it naturally serves as the measure for soliciting human input as needed.

The third issue to be addressed is to close the loop, where the human inputs should feed back to the computational methods to improve them. From a learning viewpoint, online learning and, more broadly, the notion of a life-long learning system applies here. Some studies have researched these three issues in the context of collaborative active learning in crowdsourcing for visual content analysis.[30,31]

Distribution and Systems

With big data, multimedia distribution and delivery can exploit the "intelligence" of content and users, so new systems technologies are appearing for multimedia systems.

Data-Driven Edge-Network Multimedia Distribution

Recent studies have found that edge-cloud resources can improve the performance of social media delivery compared to traditional content delivery.[32] Meanwhile, as multimedia content is generated by social network users and personal devices, data is also stored and processed at the edge of the network. Distribution over the edge network can naturally meet the demand of processing the data at different geo-distributed edge datacenters.

In recent years, online social network has greatly changed content delivery—that is, the

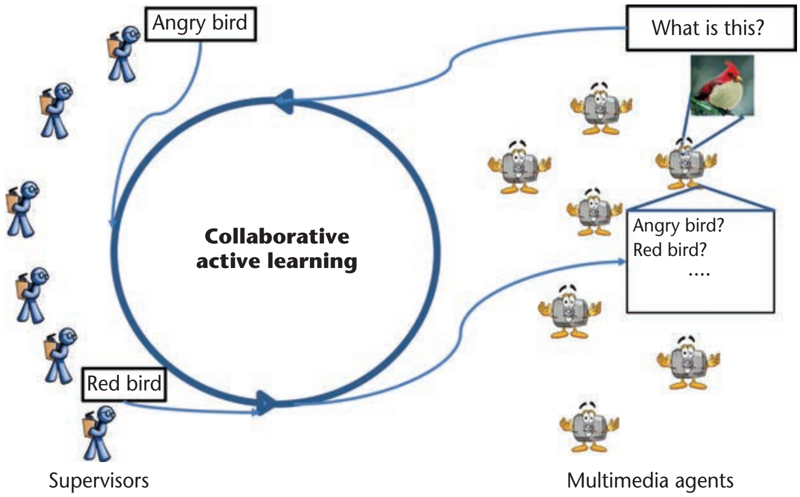


Figure 5. Illustration of a human in the loop multimedia big data analysis system exploiting crowds.

distribution of social contents has shifted from a "central-edge" manner to an "edge-edge" manner. Eytan Bakshy and his colleagues studied the social influence of people in the online social network, observing that some users can be very influential in social propagation.[33] Haitao Li, Haiyang Wang, and Jiangchuan Liu studied the content sharing in an online social network and observed the skewed popularity distribution of content and the *power-law* activity of users.[34] Giovanni Comarela and his colleagues investigated response time of social contents using collected traces and observed factors that affect the response time in social propagation.

As online social networks are affecting *dissemination* for all types of online contents, conventional content delivery paradigms need improvement using social information. Josep Pujol and his colleagues designed a social partition and replication middleware where users' friends' data can be co-located in datacenter servers.[35] Salvatore Scellato and his colleagues investigated using social *cascading* information for content delivery over the edge networks.[36] The possibility of inferring social propagation according to users' social profiles and behaviors has also been investigated,[32] allocating network resources at edge-cloud servers based on propagation predictions (see Figure 6).

Using a data-driven approach, a new trend is to exploit machine-learning techniques to learn user and content intelligence[37]—for example, in the form of mining the user's quality of experience, which in turn can make multimedia distribution more efficient.

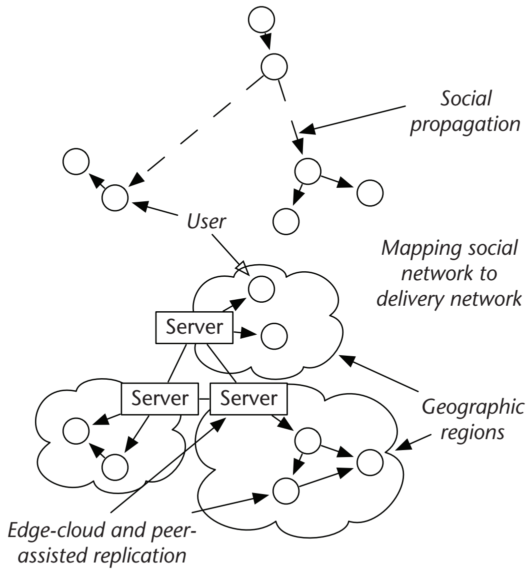


Figure 6. Edge-network social multimedia content distribution.

Streamed and In-Memory Multimedia Processing

Parallel data processing has been the mainstream of designing efficient data-processing platforms so that data could be processed in a distributed and parallel manner, improving the throughput of data processing. MapReduce (http://mapreduce.sandia.gov) is the most representative paradigm. Today's multimedia data is streamed in nature—that is, the data is generated and updated over time. An effective multimedia data-processing system must process the data in a streamed manner. Storm (https://storm.apache.org) is one such system. It creates

In-memory multimedia data processing is critical for real-time multimedia applications.

a topology to form several "pipes" for data to pass through for processing along the way.

Another trend in multimedia big data is "in-memory" processing—that is, the data is processed in the memory instead of on hard disks, significantly reducing the processing latency. In general data processing, several in-memory paradigms have been invented, such as Berkeley's Spark (https://spark.apache.org). These types of systems often implement a "cache" strategy, which can move data from hard disks to memory for repeated actions, to reduce the cost of accessing hard drivers. In the context of multimedia data processing, in-memory image and video processing is in demand.

The approaches for multimedia big data computing have a broad range of application scenarios, such as healthcare and medical applications, social media, satellite imaging, IoT, and smart cities. Future research will focus on the scale and complexity problems encountered in multimedia big data computing. For analytics, we need effective and efficient algorithms that address issues of scale and complexity. Taking a data-driven approach, such as with deep learning, will still be effective for multimedia big data analytics. However, with new developments in artificial intelligence and human-computer interaction, we see potential in systematically integrating knowledge- and data-driven approaches. For example, jointly considering deep learning and wisdom of the crowd (crowdsourcing) will be a promising future direction. In terms of systems, determining how to jointly optimize computing, storage, and communication/networking will require further research. **MM**

Acknowledgments

We thank Daixin Wang and Shaowei Liu from Tsinghua University for their contributions on representation learning and suggestions.

References

1. W. Wolf, "Cyber-Physical Systems," *Computer*, vol. 42, no. 3, 2009, pp. 88–89; doi: 10.1109/MC.2009.81.
2. A. Jacobs, "The Pathologies of Big Data," *Comm. ACM*, vol. 52, no. 8, 2009, pp. 36–44.
3. R. Ali et al., "Social Sensing: When Users Become Monitors," *Proc. 19th ACM SIGSOFT Symp. 13th European Conf. Foundations of Software Eng.*, 2011, pp. 476–479.
4. A. Madan et al., "Social Sensing for Epidemiological Behavior Change," *Proc. 12th ACM Int'l Conf. Ubiquitous Computing*, 2010, pp. 291–300.
5. H.T. Kung, "Big Data and Compressive Sensing," *Contemporary Computing*, Springer, 2012, p. 6.
6. M.K Jeong et al., "Wavelet-Based Data Reduction Techniques for Process Fault Detection," *Technometrics*, vol. 48, no. 1, 2006, pp. 26–40.
7. L.-Y. Duan et al., "Compact Descriptors for Visual Search," *IEEE MultiMedia*, vol. 21, no. 3, 2014, pp. 30–40.
8. N. Srivastava and R. Salakhutdinov, "Multimodal Learning with Deep Boltzmann Machines," *Proc. Neural Information Processing Systems Conf.* (NIPS), 2012, pp. 2231–2239.
9. F. Feng, X. Wang, and R. Li, "Cross-Modal Retrieval with Correspondence Autoencoder," *ACM Int'l Conf. Multimedia*, 2014, pp. 7–16.
10. Q. Zhang et al., "Can Visual Recognition Benefit from Auxiliary Information in Training?" *Computer Vision—ACCV* 2014, Lecture Notes in Computer Science, vol. 9003, 2015, pp 65–80.
11. Z. Niu et al., "Semi-Supervised Relational Topic Model for Weakly Annotated Image Recognition in Social Media," *IEEE Conf. Computer Vision and Pattern Recognition* (CVPR), 2014, pp. 4233–4240.
12. F. Qiu and J. Cho, "Automatic Identification of User Interest for Personalized Search," *Proc. 15th Int'l Conf. World Wide Web*, 2006, pp. 727–736.
13. E. Agichtein et al., "Learning User Interaction Models for Predicting Web Search Result Preferences," *Proc. 29th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval*, 2006, pp. 3–10.
14. J. Teevan, S.T. Dumais, and E. Horvitz, "Personalizing Search via Automated Analysis of Interests and Activities," *Proc. 28th Ann. Int'l ACM SIGIR Conference on Research and Development in Information Retrieval*, 2005, pp. 449–456.
15. M. Jiang et al., "Social Recommendation Across Multiple Relational Domains," *Proc. 21st ACM Int'l Conf. Information and Knowledge Management*, 2012, pp. 1422–1431.

16. P. Cui et al., "Who Should Share What? Item-Level Social Influence Prediction for Users and Posts Ranking," *Proc. ACM Int'l Conf. Information Retrieval* (SIGIR), 2011, pp. 185–194.
17. M. Lipczak, M. Trevisiol, and A. Jaimes, "Analyzing Favorite Behavior in Flickr," *Advances in Multimedia Modeling*, Springer, 2013, pp. 535–545.
18. X. Xie et al., "Learning User Interest for Image Browsing on Small-Form-Factor Devices," *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 2005, pp. 671–680.
19. Y. Yang et al., "User Interest and Social Influence Based Emotion Prediction for Individuals," *Proc. 21st ACM Int'l Conf. Multimedia*, 2013, pp. 785–788.
20. P. André et al., "Designing Novel Image Search Interfaces by Understanding Unique Characteristics and Usage," *Human-Computer Interaction—INTERACT*, Springer, 2009, pp. 340–353.
21. P.A. Chirita et al., "Using ODP Metadata to Personalize Search," *Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval*, 2005, pp. 178–185.
22. A. Sieg, B. Mobasher, and R. Burke, "Web Search Personalization with Ontological User Profiles," *Proc.16th ACM Conf. Information and Knowledge Management*, 2007, pp. 525–534.
23. K. Lerman, A. Plangprasopchok, and C. Wong, "Personalizing Image Search Results on Flickr," *Intelligent Information Personalization*, 2007.
24. K. Stefanidis, E. Pitoura, and P. Vassiliadis, "Managing Contextual Preferences," *Information Systems*, vol. 36, no. 8, 2011, pp. 1158–1180.
25. B. Sarwar et al., "Item-Based Collaborative Filtering Recommendation Algorithms," *Proc. ACM Int'l Conf. World Wide Web* (WWW), 2001, pp. 285–295.
26. Y. Koren, R. Bell, and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," *Computer*, vol. 42, no. 8, 2009, pp. 30–37.
27. J. Leskovec, A. Singh, and J. Kleinberg, "Patterns of Influence in a Recommendation Network," *Advances in Knowledge Discovery and Data Mining*, Springer, 2006, pp. 380–389.
28. M. Jamali and M. Ester, "Trustwalker: A Random Walk Model for Combining Trust-Based and Item-Based Recommendation," *Proc. 15th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining*, 2009, pp. 397–406.
29. S. Branson, P. Perona, and S. Belongie, "Strong Supervision from Weak Annotation: Interactive Training of Deformable Part Models," *Proc. IEEE Int'l Conf. Computer Vision* (ICCV), 2011, pp. 1832–1839.
30. G. Hua et al., "Collaborative Active Learning of a Kernel Machine Ensemble for Recognition," *Proc. IEEE Int'l Conf. Computer Vision* (ICCV), 2013, pp. 1209–1216.
31. C. Long, G. Hua, and A. Kapoor, "Active Visual Recognition with Expertise Estimation in Crowdsourcing," *Proc. IEEE Int'l Conf. Computer Vision* (ICCV), 2013, pp. 3000–3007.
32. Z. Wang et al., "Propagation-Based Social-Aware Replication for Social Video Contents," *Proc. ACM Int'l Conf. Multimedia*, 2012, pp. 29–38.
33. E. Bakshy et al., "Everyone's an Influencer: Quantifying Influence on Twitter," *Proc. ACM Int'l Con. Web Search and Data Mining* (WSDM), 2011, pp. 65–74.
34. H. Li, H. Wang, and J. Liu, "Video Sharing in Online Social Network: Measurement and Analysis," *Proc. ACM Network and Operating System Support on Digital Audio and Video Workshop* (NOSSDAV), 2012, pp. 83–88.
35. J.M. Pujol et al., "The Little Engine(s) that Could: Scaling Online Social Networks," *ACM SIGCOMM Computer Comm. Rev.*, vol. 40, no. 4, 2010, pp. 375–386.
36. S. Scellato et al., "Track Globally, Deliver Locally: Improving Content Delivery Networks by Tracking Geographic Social Cascades," *Proc. 20th Int'l Conf. World Wide Web*, 2011, pp. 457–466.
37. A. Balachandran et al., "Developing a Predictive Model of Quality of Experience for Internet Video," *Proc. ACM SIGCOMM 2013 Conf.*, pp. 339–350.

Wenwu Zhu is a full professor at Tsinghua University, China. Contact him at wwzhu@tsinghua.edu.cn.

Peng Cui is an assistant professor at Tsinghua University, China. Contact him at cuip@tsinghua.edu.cn.

Zhi Wang is an assistant professor at Tsinghua University, China. Contact him at wangzhi@sz.tsinghua.edu.cn.

Gang Hua is an associate professor at Steven Institute of Technology. Contact him at ganghua@gmail.com.

This article originally appeared in IEEE MultiMedia, *vol. 22, no. 3, 2015.*

cn *Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.*

MANAGER. Job location: Miami, FL & any other unanticipated locations in U.S. Travel Required. Duties: Resp. for full cycle software implem. delivery. Manage local & remote cross functional teams using Agile & Waterfall software methodologies to ensure delivery. Create & analyze business ROI based on software implems. for C-level executives. Conduct client interviews & workshops to understand & recommend best practices. Develop & extend enterprise processes & workflow using Enterprise HR appls such as Oracle, Microsoft, Kronos, Infor, & Ceridian. Perform enterprise solution implem. & architecture. Requires: M.S. degree in Comp. Sci, Eng. Or related field & 3 yrs. exp. in job offered or 3 yrs. exp. as a Consultant, Architect or Developer. Will accept B.S. (or foreign equiv.) & 5 yrs. exp. in the comp. ind. in lieu of M.S. & 3 yrs. exp. Concurrent exp. must incl.: 3 yrs. exp. with enterprise solution implem. & architecture & 2 yrs. exp. using Agile & Waterfall. Send resume (no calls) to: Michelle Ramirez, The Hackett Group, Inc., 1001 Brickell Bay Dr., Suite 3000, Miami, FL 33131.

BAYLOR UNIVERSITY 1845

Faculty Position

The Electrical and Computer Engineering Department of Baylor University seeks faculty applicants for a tenured/tenure-track Faculty Position at any level. Any area of expertise will be considered but applicants in computer engineering will be given special consideration. Applicants for assistant professor must demonstrate potential for sustained, funded scholarship and excellent teaching; applicants for associate or full professor must present evidence of achievement in research and teaching commensurate with the desired rank. The ECE department offers B.S., M.S., M.E. and Ph.D. degrees and is rapidly expanding its faculty size. Facilities include the *Baylor Research and Innovation Collaborative* (BRIC), a newly-established research park minutes from the main campus.

Chartered in 1845 by the Republic of Texas, Baylor University is the oldest university in Texas. Baylor has an enrollment of over 15,000 students and is a member of the Big XII Conference. Baylor's mission is to educate men and women for worldwide leadership and service by integrating academic excellence and Christian commitment within a caring community. The department seeks to hire faculty with an active Christian faith; applicants are encouraged to read about Baylor's vision for the integration of faith and learning at www.baylor.edu/profuturis/.

Applications will be considered on a rolling basis until the January 1, 2016 deadline. Applications must include:

1. **a letter of interest that identifies the applicant's anticipated rank,**
2. **a complete CV,**
3. **a concise statement of teaching and research interests,**
4. **the names and contact information for at least four professional references.**

Additional information is available at www.ecs.baylor.edu. Send materials via email to Dr. Keith Schubert at keith_schubert@baylor.edu. Please combine all submitted material into a single pdf file.

Baylor University is affiliated with the Baptist General Convention of Texas. As an Affirmative Action/Equal Employment Opportunity employer, Baylor encourages candidates of the Christian faith who are minorities, women, veterans, and persons with disabilities to apply.

NISUM TECHNOLOGIES has multiple openings for the following positions at its office in Brea, CA. Information Systems Manager: Plan & Coordinate activities of software professionals. Technical Lead: Design, develop and test software systems. QA Engineer: Plan and conduct analysis, inspection, design, test and/or integration to assure quality of projects. Software Developer: design, develop, maintain & test software applications/systems. Analyze customer requirements & custom design systems as needed. Database Administrator: Install, maintain, administer & troubleshoot databases. Systems Analyst: Plan, coordinate, design, develop, test & implement software systems & applications. Job requires min. of M.S./foreign equiv. + exp., M.S./foreign equiv., or B.S./foreign equiv. + exp. Education/Experience requirements vary depending on position level/type. Travel/relocation required. Send resume and salary history & position applied for to: Nisum Technologies, 500 S. Kraemer Blvd., Brea, CA 92821. Attn: H.R. Manager.

SENIOR SYSTEMS ADMINISTRATOR F/T (Fishkill, NY) Position involves travel to various unanticipated worksites up to 100% of the time anywhere in the United States. Must have Master's deg or the foreign equiv in Comp Sci & Software Engg, Electrical Engg, Info Systems, Electrical & Comp Engg, or related with one (1) year of exp providing support using VBlock 300/700 and best practices implementation. Monitor the entire VBlock hardware infrastructure by consolidating the UIM Provisioning center and Operations. Configuring and managing VMware on Dell, HP and Cisco servers. Configure and manage Horizon VIEW (VDI) environment. Perform Performance tuning of ESX/ESXi server, Provide support ESXi hosts and VCenter - including VMotions, Storage Vmotions, storage management and ESX/ESXi hosts upgrade. VCenter upgrade. Linux scripting providing in utilizing Virtual Center, Perform networking activities on Nexus 5K/9K and also Nexus 1000V. Provide leadership in recommending and implementing continuous process improvement, education and training requirements to management staff. Send resume: Novisync, Inc., Recruiting, 300 Westage Bus Ctr Dr, Ste 350, Fishkill, NY 12524.

 2376-113X/15/$31.00 © 2015 IEEE

CLASSIFIED LINE AD SUBMISSION DETAILS: Rates are $425.00 per column inch ($640 minimum). Eight lines per column inch and average five typeset words per line. Send copy at least one month prior to publication date to: Debbie Sims, Classified Advertising, *Computer* Magazine, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720; (714) 816-2138; fax (714) 821-4010. Email: dsims@computer.org.

In order to conform to the Age Discrimination in Employment Act and to discourage age discrimination, *Computer* may reject any advertisement containing any of these phrases or similar ones: "…recent college grads…," "…1–4 years maximum experience…," "…up to 5 years experience," or "…10 years maximum experience." *Computer* reserves the right to append to any advertisement without specific notice to the advertiser. Experience ranges are suggested minimum requirements, not maximums. *Computer* assumes that since advertisers have been notified of this policy in advance, they agree that any experience requirements, whether stated as ranges or otherwise, will be construed by the reader as minimum requirements only. *Computer* encourages employers to offer salaries that are competitive, but occasionally a salary may be offered that is significantly below currently acceptable levels. In such cases the reader may wish to inquire of the employer whether extenuating circumstances apply.

Cisco Systems, Inc. is accepting resumes for the following positions:

GLENDALE, CA: Network Consulting Engineer (Ref.# GLE3): Responsible for the support and delivery of Advanced Services to company's major accounts. Telecommuting permitted and travel may be required to various unanticipated locations throughout the United States.

LAWRENCEVILLE, GA: Hardware Engineer (Ref.# LV1): Responsible for the specification, design, development, test, enhancement, and sustaining of networking hardware. Network Consulting Engineer (Ref.# LV12): Responsible for the support and delivery of Advanced Services to company's major accounts.

RESEARCH TRIANGLE PARK, NC: Customer Support Engineer (Ref.# RTP1): Responsible for providing technical support regarding the company's proprietary systems and software. Technical Leader (Ref.# RTP 5): Lead engineering groups on projects to design, develop or test hardware or software products. Manager, Technical Support (Ref.# RTP452): Responsible for leading a team in the delivery of world-class customer support on a line of products or for a targeted group of customers. Technical Leader Services (Ref.# RTP715): Independently solve problems in broad, complex, and unique networks in Service Provider, Enterprise, and Data Center environments. Consulting Systems Engineer (Ref.# RTP30): Provide specific end-to-end solutions and architecture consulting, technical and sales support for major account opportunities at the theater, area, or operation level.

RICHARDSON, TX: Software/QA Engineer (Ref.# RIC30): Debug software products through the use of systematic tests to develop, apply, and maintain quality standards for company products. Software Engineer (Ref.# RIC3): Responsible for the definition, design, development, test, debugging, release, enhancement or maintenance of networking software.

RICHFIELD, OH: Systems Engineer (Ref.# RICH5): Provide business-level guidance to the account team or operation on technology trends and competitive threats, both at a technical and business level.

SAN FRANCISCO, CA: Inside Systems Engineer (CNG Staff) (Ref.# SF18): Responsible for conducting online product demonstrations, answering technical questions, contributing to proposals and analyzing client needs and develop technical solutions in a pre-sales capacity. Travel may be required to various unanticipated locations throughout the United States. CNG Member Technical Staff (Ref.# SF86): Architect and develop requirements documents for new products. Oversee JDM (Joint Design Manufacturer) partners through schematic, layout, bring-up, and testing. Product Manager (Ref.# SF51): Create high level marketing strategies and concepts for company solutions for markets and segments worldwide. CNG Member Technical Staff (Ref.# SF10): Design, implement, and test software for a web application used by our customers for IT management. CNG Management (Ref.# SF52): Develop sourcing strategies for key silicon technologies and unique components that are most integral for product development success.

SAN JOSE/MILPITAS/SANTA CLARA, CA: IT Engineer (Ref.# SJ7): Responsible for development, support and implementation of major system functionality of company's proprietary networking products. Network Consulting Engineer (Ref.# SJ107): Responsible for the support and delivery of Advanced Services to company's major accounts. Travel may be required to various unanticipated locations throughout the United States. User Experience Designer (Ref.# SJ587): Identify user interaction requirements and develop user experience interface specifications and guidelines. Hardware Engineer (Ref.# SJ5): Responsible for the specification, design, development, test, enhancement, and sustaining of networking hardware. Technical Marketing Engineer (Ref.# SJ178): Responsible for enlarging company's market and increasing revenue by marketing, supporting, and promoting company's technology to customers. Travel may be required to various unanticipated locations throughout the United States. Technical Marketing Engineer (Ref.# SJ15): Responsible for enlarging company's market and increasing revenue by marketing, supporting, and promoting company's technology to customers. Network Engineer (Ref.# SJ57): Responsible for the operational support of internal network systems. Diagnostic Engineer (Ref.# SJ168): Design and develop diagnostic software for verification and validation in engineering and manufacturing. Business Systems Analyst (Ref.# SJ518): Optimize operational efficiency and develop systemic process solutions. Consulting Systems Engineer (Ref.# SJ812): Provide specific end-to-end solutions and architecture consulting, technical and sales support for major account opportunities at the theater, area, or operation level. Telecommuting permitted and travel may be required to various unanticipated locations throughout the United States. Principal Engineer (Ref.# SJ673): Lead the software architecture design and development of subsystems. Contribute to functional specifications and product roadmap. Engineering Architect (Ref.# SJ851): Design, develop and deliver OpenStack Cloud solutions and reference architecture for Cisco's UCS platform. Systems Engineer (Ref.# SJ13): Provide business-level guidance to the account team or operation on technology trends and competitive threats, both at a technical and business level. Product Manager (Ref.# SJ17): Create high level marketing strategies and concepts for company solutions for markets and segments worldwide. Systems Engineer (Ref.# SJ143): Provide business-level guidance to the account team or operation on technology trends and competitive threats, both at a technical and business level. Telecommuting permitted. Senior Engineer, Mechanical (Ref.# SJ752): Provides mechanical support to engineering teams. Engineering Architect (Ref.# SJ189): Work on the cutting edge of a wide range of innovative company's WebEx uses cases. Engineering Manager (Ref.# SJ914): Schedule and evaluate the resources required for multiple projects in terms of human resources and hardware equipment allocation. Mechanical Engineer (Ref.# SJ75): Provide mechanical support to engineering teams. Hardware Engineer (Ref.# SJ558): Participate on development of Application Specific Integrated Circuit (ASIC) for next generation data center switch product family, with emphasis on routing/switching protocols. Practice Architect (Ref.# SJ759): Lead technical design process and Identify and insert the right technical resources when needed.

SEATTLE, WA: Software Engineer (Ref.# SEA1): Responsible for the definition, design, development, test, debugging, release, enhancement or maintenance of networking software.

PLEASE MAIL RESUMES WITH REFERENCE NUMBER TO CISCO SYSTEMS, INC., ATTN: M51H, 170 W. Tasman Drive, Mail Stop: SJC 5/1/4, San Jose, CA 95134. No phone calls please. Must be legally authorized to work in the U.S. without sponsorship. EOE.

www.cisco.com

Microsoft

Microsoft Corporation currently has the following openings (job opportunities available at all levels, including Principal, Senior and Lead levels):

REDMOND, WA

Artists, Art Leads and Animators: Responsible for designing and creating art assets that meet or exceed industry standards for quality while supporting Microsoft Game Studio (MGS) business goals. http://bit.ly/MSJobs_Art

Applied Scientist: Utilize knowledge in applied statistics and mathematics to handle large amounts of data using various tools. http://bit.ly/MSJobs_Data_Applied_Science

Business Managers and Business Development Managers/Business Development and Strategy Analyst Manager: Develop business opportunities for sales of software and services. http://bit.ly/MSJobs_Business_Development

Business Process Manager: Responsible for the design, implementation, and release of programs or projects. http://bit.ly/MSJobs_Fin_Plan_Analy_Contr

Consultants: Deliver design, planning, and implementation services that provide IT solutions to customers and partners. Roving Employee—requires travel up to 100% with work to be performed at various unknown worksites throughout the U.S. http://bit.ly/MSJobs_Technical_Delivery

Data Scientist: Manipulate large volumes of data, create new and improved techniques and/or solutions for data collection, management and usage. http://bit.ly/MSJobs_Data_Applied_Science

Design Verification/Validation Engineers: Responsible for ensuring the quality of Microsoft hardware products. http://bit.ly/MSJobs_Hardware_Design_Verification_Eng

Designers: Develop user interface and user interaction designs, prototypes and/or concepts for business productivity, entertainment or other software or hardware applications. http://bit.ly/MSJobs_Design

Evangelists/Technical Evangelists: Collaborate with sales teams to understand customer requirements, promote the sale of products, and provide sales support. http://bit.ly/MSJobs_Tech_Evangelist

Game/Systems Designer: Create design documents for multiple major features on large projects and the entire design on smaller projects, ensuring consistency of design. http://bit.ly/MSJobs_Game_Design

Hardware Dev., Test or Design Engineers, Hardware Engineers, Electrical Engineers, Design Engineers (all levels, including Leads and Managers): Design, implement and test computer hardware. (http://bit.ly/MSJobs_Hardware_Dev_Eng)(http://bit.ly/MSJobs_Electrical_Eng)

Hardware Dev., Test or Design Engineers, Hardware Engineers, Electrical Engineers, Design Engineers (all levels, including Leads and Managers): Design, implement and test computer hardware. Requires domestic and international travel up to 25%. (http://bit.ly/MSJobs_Hardware_Dev_Eng) (http://bit.ly/MSJobs_Electrical_Eng)

International Project Engineers/Managers: Plan, initiate, and manage information technology (IT) projects that enable the delivery of products/services that meet customer expectations and needs in each international market. http://bit.ly/MSJobs_Intl_Proj_Eng

Premier Field Engineers: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Roving Employee—requires travel up to 100% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. http://bit.ly/MSJobs_Support_Delivery

Researchers/Scientists: Conduct research and lead research collaborations that yield new insights, theories, analyses, data, algorithms, and prototypes. http://bit.ly/MSJobs_Research

Service Engineers/Managers, Service Operations Engineers, and Systems/Operations Engineers/Site Reliability Engineer: Research, design, develop, and test operating systems-level software, compilers, and network distribution software. (http://bit.ly/MSJobs_Service_Engineering) (http://bit.ly/MSJobs_IT_Serv_Eng) (http://bit.ly/MSJobs_IT_Serv_Ops)

Solution Managers: Identify and analyze internal client and partner business needs, and translate needs into business requirements and value-added solutions and solution roadmaps. http://bit.ly/MSJobs_IT_Solution_Mgmt

Support Engineers / Escalation Engineers: Install, configure, support and troubleshoot issues related to Microsoft technologies. http://bit.ly/MSJobs_Support_Eng

Build Engineer II: Responsible for delivering daily Software and Firmware builds and for the tools and automation to make the process efficient and more reliable. https://jobs-microsoft.icims.com/jobs/3722/go/job

Business Analyst (Business Strategy Analyst): Plan, forecast and develop deep business insight into revenue streams for the business. Requires Domestic and International travel up to 25%. https://jobs-microsoft.icims.com/jobs/3656/go/job

Business Analytics Specialist: Plan, develop, and deliver deep business insight into people priorities, HR strategies, and HR initiatives for the business. https://jobs-microsoft.icims.com/jobs/3432/job

Business Analytics Specialist: Plan, forecast and develop deep business insight into revenue streams for the business. https://jobs-microsoft.icims.com/jobs/3648/job

Consultant: Deliver design, planning, and implementation services that provide IT solutions to customers and partners. Roving Employee - Requires travel up to 100% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3495/job

Data and Applied Scientist: Utilize knowledge in applied statistics and mathematics to handle large amounts of data using various tools. https://jobs-microsoft.icims.com/jobs/3646/go/job

Design Verification Engineer 3: Responsible for ensuring the quality of Microsoft hardware products. Requires domestic and international travel up to 25%. https://jobs-microsoft.icims.com/jobs/3542/go/job

Designer 2: Design for the Xbox core product, accessories, incubation, and vision hardware programs. https://jobs-microsoft.icims.com/jobs/3937/go/job

Electrical Engineer 2: Lead a project team of Hardware Engineers and create the next generation of Surface Devices. Requires domestic and international travel up to 25%. https://jobs-microsoft.icims.com/jobs/3660/go/job

GM, Hardware: Lead the overall Strategy and Planning Team for the Surface organization within Microsoft. Requires domestic and international travel up to 25%. https://jobs-microsoft.icims.com/jobs/3651/job

Hardware Engineer (Senior Manufacturing Process Engineer): Design, implement and test computer hardware products that add strategic value to the company. Domestic and International travel required up to 50%. https://jobs-microsoft.icims.com/jobs/3485/go/job

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 50% with work to be performed at various unknown worksites throughout the U.S. https://jobs-microsoft.icims.com/jobs/3434/job

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 75% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3578/job

Premier Field Engineer (Premier Field Engineer Devices): Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires International travel up to 25%. https://jobs-microsoft.icims.com/jobs/3667/go/job

Security Software Engineer: Write analysis reports and produce anti-malware signatures and incident response solutions for specific threats. (https://jobs-microsoft.icims.com/jobs/3885/job) (https://jobs-microsoft.icims.com/jobs/3384/go/job)

Security Software Engineer II: Responsible for developing and testing cloud services and big data solutions. https://jobs-microsoft.icims.com/jobs/3886/job

Senior Business Development Analyst: Develop business opportunities for sales of software and services. Requires travel up to 25% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3484/job

Senior Director - IOT BD: Lead cross-organization orchestration program for IoT (Internet of Things) pioneer customers, including program processes, budgeting, executive reviews. Requires Domestic and International travel up to 25%. https://jobs-microsoft.icims.com/jobs/3671/go/job

Senior Director, Sales and Operations: Manage the phones sales and operations planning, supply planning and factory planning business-critical function of Microsoft. Requires Domestic and International travel up to 25%. https://jobs-microsoft.icims.com/jobs/3613/go/job

Senior Strategy Planner: Drive strategic initiatives and business transformation projects with implications to the direction and operational mode of a unit or whole of the devices group. https://jobs-microsoft.icims.com/jobs/3481/go/job

Senior Technical Evangelist: Collaborate with sales teams to understand customer requirements, promote the sale of products, and provide sales support. Requires Domestic and International travel up to 25%. https://jobs-microsoft.icims.com/jobs/3650/go/job

Software Engineer II: Responsible for developing or testing computer software applications, systems or services. Telecommuting Permitted. Position allows employee to reside anywhere in the U.S. and telecommute to perform work exclusively from home. https://jobs-microsoft.icims.com/jobs/3935/go/job

Sourcing Manager, Optics: Support the Microsoft global procurement process by managing the development, implementation, and alignment of global sourcing strategies and global supplier network capabilities for key commodities and spend categories to achieve business objectives. Requires Domestic and International travel up to 50%. https://jobs-microsoft.icims.com/jobs/3942/job

Support Engineer: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. Requires domestic travel up to 25%. https://jobs-microsoft.icims.com/jobs/3718/job

Senior Business Planner: Develop and implement business strategies, business models, SKUs, and pricing for a set of Microsoft networking products and services. https://jobs-microsoft.icims.com/jobs/3510/job

RENO, NV

Partners OPS Manager Launch: Plan, initiate, and manage information technology (IT) projects. Lead and guide the work of technical staff. https://jobs-microsoft.icims.com/jobs/3665/job

MOUNTAIN VIEW, CA

Data Scientist: Manipulate large volumes of data, create new and improved techniques and/or solutions for data collection, management and usage. http://bit.ly/MSJobs_Data_Applied_Science

Design Researchers: Develop user interface and user interaction designs, prototypes and/or concepts for business productivity, entertainment or other software or hardware applications. http://bit.ly/MSJobs_Design_Research

Design Verification/Validation Engineers: Responsible for ensuring the quality of Microsoft hardware products. http://bit.ly/MSJobs_Hardware_Design_Verification_Eng

Hardware Dev., Test or Design Engineers, Hardware Engineers, Electrical Engineers, Design Engineers (all levels, including Leads and Managers): Design, implement and test computer hardware. (http://bit.ly/MSJobs_Hardware_Dev_Eng)(http://bit.ly/MSJobs_Electrical_Eng)

Researchers/Scientists: Conduct research and lead research collaborations that yield new insights, theories, analyses, data, algorithms, and prototypes. http://bit.ly/MSJobs_Research

Service Engineers/Managers, Service Operations Engineers, and Systems/Operations Engineers/Site

Reliability Engineer: Research, design, develop, and test operating systems-level software, compilers, and network distribution software. (http://bit.ly/MSJobs_Service_Engineering) (http://bit.ly/MSJobs_IT_Serv_Eng) (http://bit.ly/MSJobs_IT_Serv_Ops)

SAN FRANCISCO, CA

Technical Account Managers: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. http://bit.ly/MSJobs_Delivery_Relationship_Mgmt

Designers: Develop user interface and user interaction designs, prototypes and/or concepts for business productivity, entertainment or other software or hardware applications. http://bit.ly/MSJobs_Design

Designer 2: Develop user interface and user interaction designs, prototypes and/or concepts for business productivity, entertainment or other software or hardware applications. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3633/go/job

Licensing Sales Specialist: Responsible for contributing to the overall account plan by developing appropriate licensing strategies to further grow revenue and annuity penetration. https://jobs-microsoft.icims.com/jobs/3666/go/job

Premier Field Engineer, Active Directory: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires domestic travel up to 25%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3476/job

Software Engineer II: Responsible for developing or testing computer software applications, systems or services. Requires domestic travel up to 50%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3617/job

Technical Account Manager: Assure productive use of Microsoft technologies, focusing on delivering quality through planning and governance. Requires travel up to 50% with work to be performed at various unknown worksites throughout the San Francisco/San Jose-Sunnyvale-Santa Clara metro areas. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3386/go/job

SUNNYVALE, CA

Applied Scientist: Utilize knowledge in applied statistics and mathematics to handle large amounts of data using various tools. http://bit.ly/MSJobs_Data_Applied_Science

Data Scientist: Manipulate large volumes of data, create new and improved techniques and/or solutions for data collection, management and usage. http://bit.ly/MSJobs_Data_Applied_Science

Designers: Develop user interface and user interaction designs, prototypes and/or concepts for business productivity, entertainment or other software or hardware applications. http://bit.ly/MSJobs_Design

Evangelists/Technical Evangelists: Collaborate with sales teams to understand customer requirements, promote the sale of products, and provide sales support. http://bit.ly/MSJobs_Tech_Evangelist

Solution Managers: Identify and analyze internal client and partner business needs, and translate needs into business requirements and value-added solutions and solution roadmaps. http://bit.ly/MSJobs_IT_Solution_Mgmt

IT Solution Architect: Identify and analyze internal client and partner business needs, and translate needs into business requirements and value-added solutions and solution roadmaps. https://jobs-microsoft.icims.com/jobs/3697/go/job

Multiple job openings are available for each of these categories. To view detailed job descriptions and minimum requirements, and to apply, visit the website address listed. EOE.

Microsoft Corporation currently has the following openings (job opportunities available at all levels, including Principal, Senior and Lead levels):

SAN DIEGO, CA

Hardware Dev., Test or Design Engineers, Hardware Engineers, Electrical Engineers, Design Engineers (all levels, including Leads and Managers): Design, implement and test computer hardware. (http://bit.ly/MSJobs_Hardware_Dev_Eng) (http://bit.ly/MSJobs_Electrical_Eng)

Researchers/Scientists: Conduct research and lead research collaborations that yield new insights, theories, analyses, data, algorithms, and prototypes. http://bit.ly/MSJobs_Research

Service Engineers/Managers, Service Operations Engineers, and Systems/Operations Engineers/Site Reliability Engineer: Research, design, develop, and test operating systems-level software, compilers, and network distribution software. (http://bit.ly/MSJobs_Service_Engineering) (http://bit.ly/MSJobs_IT_Serv_Eng) (http://bit.ly/MSJobs_IT_Serv_Ops)

FORT LAUDERDALE, FL

Business Managers and Business Development Managers/Business Development and Strategy Analyst Manager: Develop business opportunities for sales of software and services. http://bit.ly/MSJobs_Business_Development

Associate Architect: Translate business requirements into technology requirements for inclusion in contracts and statements of work. https://jobs-microsoft.icims.com/jobs/3690/job

Sales Excellence Manager: Measure and track an efficient Segment Sales Excellence function spanning scorecard, Rhythm of Business, and business management. https//jobs-microsoft.com/jobs/3741/go/job

Senior Technical Advisor: Provide technical advice and support on issues experienced with Microsoft technologies. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3496/job

Solutions Architect, LATAM: Engage w strategic customers to solve business needs by conceiving, designing, & implementing transformative technical architecture on Microsoft Azure cloud-based platform & rel. tech. Requires domestic travel up to 50%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3571/job

Technology Solutions Professional: Drive product win rates by proving the value of products to customers and partners. Requires domestic travel up to 75%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3489/job

Solutions Sales Specialist, Education: Enhance Microsoft customer relationship from capability development perspective by articulating value of services & solutions & identifying competition gaps in targeted accounts. Requires travel up to 25% with work to be performed at various unknown worksites throughout the U.S. and Latin America. https://jobs-microsoft.icims.com/jobs/3622/go/job

DOWNERS GROVE, IL

Senior Software Engineer (Senior Architect): Responsible for developing or testing computer software applications, systems or services. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3678/go/job

EDINA, MN

Solution Specialist: Enhance Microsoft customer relationship from capability development perspective by articulating value of services & solutions & identifying competition gaps in targeted accounts. Requires domestic travel up to 25% with work to be performed at various unknown worksites throughout the U.S. https://jobs-microsoft.icims.com/jobs/3519/job

Technical Account Manager: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. Requires travel up to 75% to various client sites in the Minneapolis region. Telecommuting permitted. https://jobs-microsoft.com/jobs/3744/go/job

Sales Excellence Manager: Plan, forecast and develop deep business insight into revenue streams for the business. Requires travel up to 25% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/4216/job.

CAMBRIDGE, MA

Program Managers: (all levels, including Leads and Managers) Coordinate program development of computer software applications, systems or services. (http://bit.ly/MSJobs_ProgMgr) (http://bit.ly/MSJobs_HW_ProgMgr) (http://bit.ly/MSJobs_ProdQlty_Supp) (http://bit.ly/MSJobs_IT_ProgMgr)

Data Scientist: Manipulate large volumes of data, create new and improved techniques and/or solutions for data collection, management and usage. http://bit.ly/MSJobs_Data_Applied_Science

Consultants: Deliver design, planning, and implementation services that provide IT solutions to customers and partners. Requires domestic and international travel up to 25%. http://bit.ly/MSJobs_Technical_Delivery

Service Engineers/Managers, Service Operations Engineers, and Systems/Operations Engineers/Site Reliability Engineer: Research, design, develop, and test operating systems-level software, compilers, and network distribution software. (http://bit.ly/MSJobs_Service_Engineering) (http://bit.ly/MSJobs_IT_Serv_Eng) (http://bit.ly/MSJobs_IT_Serv_Ops)

Assc. Architect: Define and align the business objectives of a customer to the strategy and solution to achieve the objectives. Domestic and International travel required up to 50%. https://jobs-microsoft.icims.com/jobs/3653/go/job

Cloud Solution Architect: Deliver design, planning and implementation services that provide IT solutions to customers and partners. Requires domestic travel up to 25%. https://jobs-microsoft.icims.com/jobs/3502/job

Consultant MCS COE: Deliver design, planning, and implementation services that provide IT solutions to customers and partners. Requires Domestic and International travel up to 75%. https://jobs-microsoft.icims.com/jobs/3693/go/job

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 50% with work to be performed at various unknown worksites throughout the U.S. https://jobs-microsoft.icims.com/jobs/3691/go/job

CHARLOTTE, NC

Support Engineers / Escalation Engineers: Install, configure, support and troubleshoot issues related to Microsoft technologies. http://bit.ly/MSJobs_Support_Eng

Technical Account Managers: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. http://bit.ly/MSJobs_Delivery_Relationship_Mgmt

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires domestic travel up to 50%. https://jobs-microsoft.icims.com/jobs/3608/job

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 50% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3380/job

CHEVY CHASE, MD

Partner Sales Executive Distribution: Enhance Microsoft customer relationship from capability development perspective by articulating value of services & solutions & identifying competition gaps in targeted accounts. Requires travel up to 25% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3567/job

WASHINGTON DC

Senior Business Strategy Manager: Analyze market and technology trends and data to develop creative solutions to key business strategy issues. https://jobs-microsoft.icims.com/jobs/3738/go/job.

NEW YORK, NY

Consultants: Deliver design, planning, and implementation services that provide IT solutions to customers and partners. Requires domestic and international travel up to 25%. http://bit.ly/MSJobs_Technical_Delivery

Account Executive: Manage IT customer account business that meets or exceeds revenue targets. Requires travel up to 50% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3607/go/job

Account Executive, Major: Build and maintain business relationships with major Microsoft customer accounts to deliver new and renewed Enterprise Agreements and net new opportunity revenue. https://jobs-microsoft.icims.com/jobs/3422/go/job

Account Manager: Grow and maintain agency/advertiser relationships that drive revenue. https://jobs-microsoft.icims.com/jobs/3575/job

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 50% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3379/job

Search Account Manager, Strategic: Build trusted-advisor relationships with highly valued premium advertisers. https://jobs-microsoft.icims.com/jobs/3878/go/job

Solution Specialist DC: Enhance Microsoft customer relationship from capability development perspective by articulating value of services & solutions & identifying competition gaps in targeted accounts. Requires travel up to 25% with work to be performed at various unanticipated worksites throughout the U.S. https://jobs-microsoft.icims.com/jobs/3513/job

IRVING, TX

Consultants: Deliver design, planning, and implementation services that provide IT solutions to customers and partners. Roving Employee—requires travel up to 100% with work to be performed at various unknown worksites throughout the U.S. http://bit.ly/MSJobs_Technical_Delivery

Support Engineers / Escalation Engineers: Install, configure, support and troubleshoot issues related to Microsoft technologies. http://bit.ly/MSJobs_Support_Eng

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 25% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3433/job

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 75% with work to be performed at various unknown worksites throughout the U.S. https://jobs-microsoft.icims.com/jobs/3573/job

IRVINE, CA

Account Technology Strategist: Identify and analyze internal client and partner business needs, and translate needs into business requirements and value-added solutions and solution roadmaps. https://jobs-microsoft.icims.com/jobs/3881/job

HOUSTON, TX

Solutions Sales Professional/Specialist: Enhance the Microsoft customer relationship from a capability development perspective by articulating the value of our services and solutions and identifying competition gaps in targeted accounts. http://bit.ly/MSJobs_Solution_Sales

Premier Field Engineers: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires travel up to 25% with work to be performed at various unknown worksites throughout the U.S. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3487/job

Technical Account Manager: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. Requires travel up to 25% throughout the U.S. http://jobs-microsoft.icims.com/jobs/3941/go/job

HUMACAO, PUERTO RICO

Principal Software Engineer, IT: Responsible for developing or testing computer software applications, systems or services. https://jobs-microsoft.icims.com/jobs/3901/job

LOS ANGELES, CA

Technical Account Managers: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. http://bit.ly/MSJobs_Delivery_Relationship_Mgmt

Technical Account Manager: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. Requires travel throughout the U.S. up to 25%. https://jobs-microsoft.icims.com/jobs/3740/go/job

Technical Account Manager: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. Requires local travel up to 75%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3618/job

Technical Account Manager: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3507/go/job

Chicago, IL

Solutions Sales Professional/Specialist: Enhance the Microsoft customer relationship from a capability development perspective by articulating the value of our services and solutions and identifying competition gaps in targeted accounts. http://bit.ly/MSJobs_Solution_Sales

Solution Specialist, Platform: Enhance Microsoft customer relationship from capability development perspective by articulating value of services & solutions & identifying competition gaps in targeted accounts. Requires domestic travel up to 25%. https://jobs-microsoft.icims.com/jobs/3686/job

ISELIN, NJ

Technical Account Manager: Assure productive use of Microsoft technologies, focusing on delivery quality through planning and governance. Requires travel up to 25% to various client sites throughout the New York City metropolitan area. https://jobs-microsoft.com/jobs/3732/go/job

Technical Evangelist: Collaborate with sales teams to understand customer requirements, promote the sale of products, and provide sales support. Requires travel up to 50% with work to be performed at various unknown worksites throughout the U.S. https://jobs-microsoft.icims.com/jobs/3658/job

Cloud Solution Architect: Architect and deploy Microsoft cloud solutions for customers. Requires travel up to 25% with work to be performed at various unknown worksites throughout the U.S. https://jobs-microsoft.icims.com/jobs/3700/go/job

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires domestic travel up to 75%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3629/job

COLUMBUS, OH

Senior Solution Architect: Responsible for the design, implementation, and release of programs or projects. Requires travel throughout the United States up to 50%. Telecommuting permitted. https://jobs-microsoft.icims.com/3733/go/job

BOISE, ID

Support Engineers / Escalation Engineers: Install, configure, support and troubleshoot issues related to Microsoft technologies. http://bit.ly/MSJobs_Support_Eng

Multiple job openings are available for each of these categories. To view detailed job descriptions and minimum requirements, and to apply, visit the website address listed. EOE.

Microsoft®

Microsoft Corporation currently has the following openings (job opportunities available at all levels, including Principal, Senior and Lead levels):

ALPHARETTA, GA

Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires domestic travel up to 50%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3585/job

BURLINGTON, MA

Senior Service Engineer (Senior Systems Engineer): Research, design, develop, and test operating systems-level software, compilers, and network distribution software. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3727/go/job

FARGO, ND

Senior Technical Advisor: Mentor support partners on technical support issues, process issues, delivery and problem solving processes. https://jobs-microsoft.icims.com/jobs/3402/go/job

MADISON, WI

Senior Software Engineer: Responsible for developing or testing computer software applications, systems or services. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3909/job

MALVERN, PA

Senior Premier Field Engineer: Provide technical support to enterprise customers, partners, internal staff or others on mission critical issues experienced with Microsoft technologies. Requires domestic travel up to 100%. Telecommuting permitted. https://jobs-microsoft.icims.com/jobs/3483/job

WILSONVILLE, OR

Hardware Dev., Test or Design Engineers, Hardware Engineers, Electrical Engineers, Design Engineers (all levels, including Leads and Managers): Design, implement and test computer hardware. (http://bit.ly/MSJobs_Hardware_Dev_Eng)(http://bit.ly/MSJobs_Electrical_Eng)

Multiple job openings are available for each of these categories. To view detailed job descriptions and minimum requirements, and to apply, visit the website address listed. EOE.

FIU | FLORIDA INTERNATIONAL UNIVERSITY

Florida International University is a comprehensive university offering 340 majors in 188 degree programs in 23 colleges and schools, with innovative bachelor's, master's and doctoral programs across all disciplines including medicine, public health, law, journalism, hospitality, and architecture. FIU is Carnegie-designated as both a research university with high research activity and a community-engaged university. Located in the heart of the dynamic south Florida urban region, our multiple campuses serve over 55,000 students, placing FIU among the ten largest universities in the nation. Our annual research expenditures in excess of $100 million and our deep commitment to engagement have made FIU the go-to solutions center for issues ranging from local to global. FIU leads the nation in granting bachelor's degrees, including in the STEM fields, to minority students and is first in awarding STEM master's degrees to Hispanics. Our students, faculty, and staff reflect Miami's diverse population, earning FIU the designation of Hispanic-Serving Institution. At FIU, we are proud to be 'Worlds Ahead'! For more information about FIU, visit fiu.edu.

The School of Computing and Information Sciences (SCIS) seeks exceptionally qualified candidates for tenure-track and tenured faculty positions at all levels as well as non-tenure track faculty positions at the level of Instructor, including visiting instructor appointments. SCIS is a rapidly growing program of excellence at the University, with 30 tenure-track faculty members and over 2,000 students, including over 80 Ph.D. students. SCIS offers B.S., M.S., and Ph.D. degrees in Computer Science, an M.S. degree in Telecommunications and Networking, an M.S. degree in Cybersecurity, and B.S., B.A., and M.S. degrees in Information Technology. SCIS has received over $22M in the last four years in external research funding, has six research centers/clusters with first-class computing and support infrastructure, and enjoys broad and dynamic industry and international partnerships.

Open-Rank Tenure Track/Tenured Positions (Job ID# 508676)

SCIS seeks exceptionally qualified candidates for tenure-track and tenured faculty positions at all levels. We seek well-qualified candidates in all areas; researchers in the areas of computer systems, cybersecurity, cognitive computing, data science, health informatics, and networking are particularly encouraged to apply. Preference will be given to candidates who will enhance or complement our existing research strengths.

Ideal candidates for junior positions should have a record of exceptional research in their early careers. Candidates for senior positions must have an active and proven record of excellence in funded research, publications, and professional service, as well as a demonstrated ability to develop and lead collaborative research projects. In addition to developing or expanding a high-quality research program, all successful applicants must be committed to excellence in teaching at both the graduate and undergraduate levels. An earned Ph.D. in Computer Science or related disciplines is required.

Non-tenure track instructor positions (Job Opening 507474)

We seek well-qualified candidates in all areas of Computer Science and Information Technology. Ideal candidates must be committed to excellence in teaching a variety of courses at the undergraduate level. A graduate degree in Computer Science or related disciplines is required; significant prior teaching and industry experience and/or a Ph.D. in Computer Science is preferred.

HOW TO APPLY:

Qualified candidates for open- rank faculty positions are encouraged to apply to (Job Opening ID #508676); and candidates for instructor positions are encouraged to apply to (Job Opening ID# 507474). Submit applications at facultycareers.fiu.edu and attach cover letter, curriculum vitae, statement of teaching philosophy, research statement, etc as *individual attachments*. Candidates will be required to provide names and contact information for at least three references who will be contacted *as determined by the search committee*. To receive full consideration, applications and required materials should be received by December 31st, 2015. Review will continue until position is filled.

If you are interested in a visiting appointment please contact the department directly by emailing Dr. Mark Weiss at Weiss@cis.fiu.edu. All other applicants should apply by going to facultycareers.fiu.edu.

FIU is a member of the State University System of Florida and an Equal Opportunity, Equal Access Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, or any other characteristic protected by law.

PROGRAMMER ANALYST: Design, develop, test & implement Web and Client/Server Technologies using knowledge in C#.Net, ASP.Net, ASP.Net MVC, WCF Services (SOAP & REST), HTML5, CSS3, XSD, XML, SQL Server 2008 R2/2012 , JavaScript , JQuery , ADO.Net, ADO.Net Entity framework ,Windows XP/VISTA/7/8 Visual Studio 2010,2012. Must be willing to travel & relocate. Requires MS in Computer science, Engineering or related. Mail resumes to Keypixel Software Solutions LLC, 777 Washington Rd, Ste 1, Parlin, NJ 08859.

SR. MANAGER. Job location: Miami, FL & any other unanticipated locations in U.S. Travel Required. Duties: Resp. for analyzing, designing, developing, testing, debugging, modifying, enhancing, implementing, integrating and documenting Business Intelligence (BI) computer software appls. using Hyperion Essbase, Hyperion financial tools, Hyperion Planning, SmartView, RDBMS & VB. Resp. for appl. design, product functionality & tech. issues & appl. prototyping/testing. Monitor project team progress & perform. & complete all follow-up projects work incl. reporting. Requires: M.S. degree in Comp. Sci, Eng. or related field & 3 yrs. exp. in job offered or 3 yrs. exp. as a Mgr. or Sr. Consultant. Will accept B.S. (or foreign equiv.) & 5 yrs. exp. in the comp. ind. in lieu of M.S. & 3 yrs. exp. Concurrent exp. must incl.: 3 yrs. exp. analyzing, designing, developing & implementing BI software appls. using Hyperion Essbase & 3 yrs. exp. using Hyperion Financial Tools. Send resume (no calls) to: Michelle Ramirez, The Hackett Group, Inc., 1001 Brickell Bay Dr., Suite 3000, Miami, FL 33131.

LEAD DBA ADMINISTRATOR F/T (Poughkeepsie, NY) Position involves travel to various unanticipated worksites up to 100% of the time anywhere in the United States. Must have Bach deg or the foreign equiv in Electronics & Communication Engg, Electrical Engg, or related w/5 yrs of progressive exp or a Master deg or the foreign equiv in Electronics & Communication Engg, Electrical Engg, or related w/1 yr exp managing & leading a team of 3 Administrators in Database Design, Implementation, SQL Tuning in DB2/Oracle environments. Provide project plan estimation and rollout strategy in collaboration with the project manager. Mentor, develop and train staff members as needed. Create and maintains databases objects including Tablespace, Table, Index, views. Planning for backup and recovery strategies. Review performance, maintenance DB2 Utilities: Unload, Load, Reorgs, Run Stats and Image copy. Provide subject matter expertise using following tools: DB2, COBOL, Oracle 10g, ERWIN, SQL Navigator, IBM Data Studio 4.1, DB2 Connect, Visual Explain, Microsoft Query, PL/SQL, Informatica PowerCenter 9.1/8.6, Java, JCL, ChangeMan, and Control-M. Send resume: Indotronix Int.l Corp., Recruiting (KN), 331 Main St, Poughkeepsie, NY 12601.

PROGRAMMER ANALYST: Design, develop, integrate, test & implement application software utilizing knowledge in Agile environment & utilizing technologies like Net Framework 4.0/4.5, C#.Net, ASP.Net Web Apps, Web Services, Ado.Net, WinForms, LINQ, Visual Studio .Net, Sql Server Management Studio, JavaScript,VB Script, T- Sql,SQL server, Oracle, HTML, XML, CSS, TFS, IIS,MS-Unit. Must be willing to travel & reloc to unanticipated client location throughout the US. Requires MS in comp sci, eng or rel. Mail resumes to Code Ace Solutions Inc. 50 Cragwood Rd, Ste 217, South Plainfield, NJ 07080.

SOFTWARE ENGINEER - design, develop, test & implement application s/w utilizing knowledge of C#, VB.Net, ASP.Net, Site Core ,HTML, CSS, Jquery, JavaScript, Knockout Js, Bootstrap, Angular Js Entity Framework, Oracle, SQL Server, SSIS, Unit Testing. Must be willing to travel and reloc. Reqs MS in comps ci, eng or rel. Mail resumes to Strategic Resources International, Inc. 777 Washington Rd, Suite 2, Parlin, NJ 08859.

TENURE-TRACK ASSISTANT PROFESSOR in Networking Department of Computer and Information Sciences Applications are invited for a tenure-track assistant professor position in Networking and associated research areas to begin Fall 2016. For application, please visit apply .interfolio.com/32786 eoe.

Microsoft

Microsoft Corporation currently has the following openings (job opportunities available at all levels, including Principal, Senior and Lead levels):

REDMOND, WA

Program Managers: (Principal and Lead ONLY) Coordinate program development of computer software applications, systems or services. Requires domestic and international travel up to 25%. (http://bit.ly/MSJobs_ProgMgr) (http://bit.ly/MSJobs_HW_ProgMgr) (http://bit.ly/MSJobs_ProdQlty_Supp)(http://bit.ly/MSJobs_IT_ProgMgr)

Program Managers: (Principal and Lead ONLY) Coordinate program development of computer software applications, systems or services. (http://bit.ly/MSJobs_ProgMgr) (http://bit.ly/MSJobs_HW_ProgMgr) (http://bit.ly/MSJobs_ProdQlty_Supp) (http://bit.ly/MSJobs_IT_ProgMgr)

Operations Program Managers: Plan, initiate, and manage information technology (IT) projects. (http://bit.ly/MSJobs_Ops_PM)

Operations Program Manager: Plan, initiate, and manage information technology (IT) projects. Lead and guide the work of technical staff. Requires travel throughout the U.S. and internationally up to 25%. https://jobs-microsoft.icims.com/jobs/3737/go/job

Operations Program Manager (Security Project Manager): Plan, initiate, and manage information technology (IT) projects. Requires domestic and international travel up to 50%. https://jobs-microsoft.icims.com/jobs/4048/go/job

MOUNTAIN VIEW, CA

Program Managers: (Pall levels, including Leads and Managers) Coordinate program development of computer software applications, systems or services. (http://bit.ly/MSJobs_ProgMgr) (http://bit.ly/MSJobs_HW_ProgMgr) (http://bit.ly/MSJobs_ProdQlty_Supp) (http://bit.ly/MSJobs_IT_ProgMgr)

SUNNYVALE, CA

Program Managers: (Pall levels, including Leads and Managers) Coordinate program development of computer software applications, systems or services. (http://bit.ly/MSJobs_ProgMgr) (http://bit.ly/MSJobs_HW_ProgMgr) (http://bit.ly/MSJobs_ProdQlty_Supp) (http://bit.ly/MSJobs_IT_ProgMgr)

FORT LAUDERDALE, FL

Operations Program Managers: Plan, initiate, and manage information technology (IT) projects. (http://bit.ly/MSJobs_Ops_PM)

Program Managers: Coordinate program development of computer software applications, systems or services. (http://bit.ly/MSJobs_ProgMgr) (http://bit.ly/MSJobs_HW_ProgMgr) (http://bit.ly/MSJobs_ProdQlty_Supp)(http://bit.ly/MSJobs_IT_ProgMgr)

Multiple job openings are available for each of these categories. To view detailed job descriptions and minimum requirements, and to apply, visit the website address listed. EOE.

Microsoft

Microsoft Corporation currently has the following openings (job opportunities available at all levels, including Principal, Senior and Lead levels):

REDMOND, WA

Software Engineers and Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. Requires domestic and international travel up to 25%. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

Research Software Development Engineers (all levels): Responsible for conducting applied research into new products and services through software engineering techniques. http://bit.ly/MSJobs_Research_Software_Engineer

MOUNTAIN VIEW, CA

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

SAN FRANCISCO, CA

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

SUNNYVALE, CA

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

Research Software Development Engineers (all levels): Responsible for conducting applied research into new products and services through software engineering techniques. http://bit.ly/MSJobs_Research_Software_Engineer

SAN DIEGO, CA

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

PALO ALTO, CA

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

CAMBRIDGE, MA

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

DURHAM, NC

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

RESTON, VA

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

NEW YORK, NY

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

Research Software Development Engineers (all levels): Responsible for conducting applied research into new products and services through software engineering techniques. http://bit.ly/MSJobs_Research_Software_Engineer

FARGO, ND

Software Engineers, Security Software Engineers, Software Development Engineers in Test (all levels, including Leads and Managers): Responsible for developing or testing computer software applications, systems or services. (http://bit.ly/MSJobs_SDE) (http://bit.ly/MSJobs_IT_SDE)

Multiple job openings are available for each of these categories. To view detailed job descriptions and minimum requirements, and to apply, visit the website address listed. EOE.

SOFTWARE ENGINEER: Design & devp sw apps for visual analytics suite of sw util high-level prog languages & tools, incl C, Java, Eclipse, & Agile devp methods. Req MS or foreign equiv in Comp Sci, or rtd field. Req educ or exp in: design & devp comp sw apps util high-level, obj-oriented prog languages, incl C, Java, Python, Eclipse, & Agile devp methods, incl Scrum; design & devp UI for SW apps using Windows; design, implement & write unit tests for sw components; troubleshoot & debug sw defects; & investigate issues rtd to sw apps, analyze code for origin of bugs, & perform troubleshoot & debug. Position at Tableau Software, Inc. in Seattle, WA. To apply, please e-mail resume to Jobstableau@tableau.com and ref Job ID: SE2.

SOFTWARE ENGINEER: Design & devp framework for testing visual analytics sw util high-level, obj-oriented prog languages, incl C#. Req BS or foreign equiv in Comp Sci, Comp Eng, or rtd field, & 2 yrs exp as SW Eng, incl exp in: design, devp & implement auto test scripts & test report using obj-oriented prog languages, incl C#; design & test highly interactive web-based apps that use HTML5, JavaScript & CSS3; design & devp test plans, test scenarios, test cases, test reports & doc for manual & auto tests; write functional, unit & integration tests for back-end apps using APIs & auto test suites; investigate prod issues, troubleshoot, analyze layers of code for origin of bugs, debug, & recommend sols to cross-functional teams; & perform data analytics using RDBMS,

DATA STRUCTURES, INCL DICTIONARIES & LINKED LISTS, & RELATIONAL DBASES. Position at Tableau Software, Inc. in Seattle, WA. To apply, please e-mail resume to Jobstableau@tableau.com and ref Job ID: SE1.

SOFTWARE ENGINEER: Design & devp customer-facing aspects of visual analytics sw prods util high level prog languages, incl C++, C#, & Python. Req MS or foreign equiv in Comp Sci, or rtd field, & 2 yrs exp in: design & devp comp sw apps util high-level, obj-oriented prog languages, incl C++, C#, & Python; perform troubleshoot & debug to solve issues; design & devp test plans, test scenarios, test cases, test reports & doc for manual & auto tests util APIs & auto test suites; perform load test, end-to-end test, & unit test of sw apps & web svcs; & util relational dbases, incl SQL, SQL Server, DDL & DML queries. Position at Tableau Software, Inc. in Kirkland, WA. To apply, please e-mail resume to Jobstableau@tableau.com and ref Job ID: SE3.

SPLUNK INC. has the following job opportunities in San Francisco, CA: **Senior Software Engineer (REQ#9PHS8E)**. Dev automation infrastructure & automated test cases. **Software Engineer (REQ#9M4VA5)**. Create fault-tolerant dist sys in multi-threaded, multi-process envir. **Senior Sales Engineer (REQ#9U43PV)**. Drive tech eval & selection stages of the sales process. Must be able to travel to unanticipated locations throughout the U.S., including frequent travel to Milwaukee & Chicago. Telecom ok.Refer to Req# & mail resume to Splunk Inc., ATTN: J. Aldax, 250 Brannan Street, San Francisco CA 94107. Individuals seeking employment at Splunk are considered without regards to race, religion, color, national origin, ancestry, sex, gender, gender identity, gender expression, sexual orientation, marital status, age, physical or mental disability or medical condition (except where physical fitness is a valid occupational qualification), genetic information, veteran status, or any other consideration made unlawful by federal, state or local laws. To review US DOL's EEO is The Law notice please visit: https://careers.jobvite.com/Splunk/EEO_poster.pdf. To review Splunk's EEO Policy Statement please visit: http://careers.jobvite.com/Careers/Splunk/EEO-Policy-Statement.pdf. Pursuant to the San Francisco Fair Chance Ordinance, we will consider for employment qualified applicants with arrest and conviction records.

EXPEDIA, INC. currently has openings for the following opportunities in our Bellevue, WA office (various/levels/types:) Software Engineers: (728.SWE-DEC) Design, implement, and debug software for computers including algorithms and data structures. Database Developers: (728.DBD-DEC) Coordinate changes to computer databases, test and implement the database applying knowledge of database management systems. Technical Product Managers: (728.TPM-DEC) Gather detailed business requirements from stakeholders and work closely with technology staff to translate requirements into functional designs and specifications. Managers, Database Development: (728.1427) Design, implement, and manage highly-optimized relational database systems. Write documentation and communicate database design. Send your resume to: Expedia Recruiting, 333 108th Avenue NE, Bellevue, WA 98004. Must reference position and Job & Job ID# listed above.

EXPEDIA, INC. currently has openings for the following opportunities in our San Francisco, CA office (various/levels/types:) Software Engineers: (728.SWE-DEC-SF) Design, implement, and debug software for computers including algorithms and data structures. Senior Oracle Applications Developers: (728.712) Drive design, development, and maintenance of systems. Provide services to Finance and Engineering organizations by supporting Oracle Financial, Discoverer Reports, Payment Plus, and other tools that are required to support ongoing finance-related day-to-day activities. Send your resume to: Expedia Recruiting, 333 108th Avenue NE, Bellevue, WA 98004. Must reference position and Job & Job ID# listed above.

HOTWIRE, INC. currently has openings for the following opportunities in our San Francisco, CA office (various/levels/types:) Software Engineers: (728.SWE-DEC-HW) Design, implement, and debug software for computers including algorithms and data structures. Send your resume to: Hotwire/Expedia Recruiting, 333 108th Avenue NE, Bellevue, WA 98004. Must reference position and Job & Job ID# listed above.

EXPEDIA, INC. currently has openings for the following opportunities in our Dallas, TX office (various/levels/types:) Software Engineers: (728.SWE-DAL) Design, implement, and debug software for computers including algorithms and data structures. Send your resume to: Expedia Recruiting, 333 108th Avenue NE, Bellevue, WA 98004. Must reference position and Job & Job ID# listed above.

CLASSIC VACATIONS, LLC. currently has openings for the following opportunities in our Dallas, TX office (various/levels/types:) Web Developers: (728.792) Contribute to front end development to enhance and design internet travel service website. Specifically, participate in walk-throughs, technical design, and all stages from implementation to release. Send your resume to: Classic Vacations/Expedia Recruiting, 333 108th Avenue NE, Bellevue, WA 98004. Must reference position and Job & Job ID# listed above.

JUNIPer NETWORKS

Juniper Networks is recruiting for our Sunnyvale, CA office:

Test Engineer #15583: Develop test plans for functional level testing of Juniper routers to enrich quality of releases for new features.

Software Engineer #32292: Work from user stories, explore detailed use cases, define UI concepts and guide development through all of the details that culminate in an excellent user experience for Juniper's firewall security products.

QA Engineer #30129: Develop & execute manual and automation tests for Juniper's management products leveraging technologies such as Python and Robot. Set-up test environments consisting of routing, switching and security platforms with associated management products.

QA Engineer Staff #36408: Perform feature testing of networking protocols for internet routing devices. Dev. functional, performance & scaling tests for high-performance secure networking products.

Software Engineer #20410: Design, implement, test and maintain software features for Juniper Networking Operating System (JUNOS).

Software Engineer #18277: Gather system requirements and translate the requirements into high quality software for Juniper Router Platform Services and maintain it.

Mail single-sided resume with job code # to
Juniper Networks
Attn: MS A.8.429A
1133 Innovation Way
Sunnyvale, CA 94089

SOFTWARE DEVELOPER – Assist in development of software for cloud based scheduling systems and desktop management suites for multiple industries. Master's degree in Computer Science, Systems Eng. or related field, or foreign equivalent. Must be proficient in C#.Net; ASP.Net; SQL Server; Java script; Visual Studio; Win Forms; Visual Basic; Ninject, Twilio, TeamCity, Resharper, Entity Framework, WCF, WPF, RFID technology. Resumes to: Elite Software Inc., 4001 W. Newberry Rd, Gainesville, FL 32607.

SR. SYSTEMS ADMIN req'd to dsgn, install, upgrade & maintain SAP BOBJ & Data Srvcs. Monitor, perform sizing, setup security /arch. diagrams/SOPs/ test plans/EU Docs. Dvlpmnt using tools Webi, Crystal, Lumira, UDT/IDT, Explorer, Dsgn Studio, DS Dsgnr, Java, VB/.NET, Oracle, DB2, Teradata. MS degree in Comp. Sci/Apps, IT/IS, Sci., Engg or closely rltd field plus 2 years of exp. in job, or alternatively, a BS degree in one of the same or closely rltd field plus 5 yrs of progressively responsible experience in job. Foreign Equiv. Degree Ok. Mail Resumes to Chadwick Riley Corp, 12443 San Jose Blvd #402 Jacksonville FL-32223.

PROGRAMMER ANALYST: Design, Develop, Implement, Validate, test and maintain the activities for Client/ Server, Web based and Oracle Applications 11i/R12 Financial Modules software. Must have knowledge in SQL, PLSQL, SQL Loader, Unix, Shell Scripting, OAF, ADF, JAVA, JSP, C#.net, ASP.net, Oracle 1g/11g, XML/ BI Publisher, Workflow Builder, J2EE, Spring, Struts, ETL, Unix and Windows XP/7/NT. Must be willing to travel & reloc to unanticuipated client locations throughout the US. Reqs MS in comp sci, eng, business or rel Mail resumes to Nitya Software Solutions Inc. 9690 South 300, Ste 319, Salt Lake City, Utah 84070.

COMPUTER PROGRAMMER: Create, modify & test the code, forms, & script that allow computer applications to run. Work from specifications drawn up by software developers. Assist software developers by analyzing user needs & designing software solutions. Creating User Experience Documents, visual design, and prototypes based on user requirements. Utilize HTML5, CSS3, NodeJS, GruntJS, Git, Adobe InDeisgn, jQuery, AngularJS. Will work in unanticipated locations. Requires 2 yrs experience. Send resume to SelSoft, Inc., 303 S. Jupiter Suite 110, Allen, TX 75002.

ERICSSON INC. has openings for:- CUSTOMER SOLUTIONS SALES MANAGER in PLANO, TX to lead and grow Cloud business by focusing on creating solutions for telecom operators, enterprises and channels. Frequent domestic travel required. Job ID: 15-TX-1397. ENGINEER-SYSTEM in EL SEGUNDO, CA to analyze technical requirements improvements, influence the product roadmap and work with developers as product owner. Job ID: 15-CA-3397. To apply please mail resume to: Ericsson Inc. 6300 Legacy Drive, R1-C12 Plano, TX 75024 and indicate appropriate Job ID.

OOYALA INC. has openings in our Santa Clara, CA location for Software Engineer (6688.33) Build interactive data visualizations using the latest web technologies, including D3 & CoffeeScript, working with terabyte-scale datasets; Associate Product Specialist (6688.75) Design & build a new tool (billing & tooling) that will optimize packages for hundreds of customers worldwide; Associate Product Specialist (6688.52) Design & build a new tool (billing & tooling) that will optimize packages for hundreds of customers worldwide, further defining Ooyala as a trusted partner; Ooyala Inc. has openings in our Plano, TX location for QA Manager (6688.80) Build & lead a team of QA Engineers & Software Developers in Test responsible for delivering world class products using agile practices in a SaaS environment & at web scale; Backend Developer (6688.76) Review requirements & translate them to high level design. Send resume to HR, 4750 Patrick Henry Drive, Santa Clara, CA 95054. Must ref. job code above when applying.

Samsung Semiconductor Inc.

has the following jobs in **San Jose, CA**:

SW Engineer [Req #5JL1613] Design & dvlpmt of control plane SW modules for a distributed storage SYS. **Test Dvlpmt Engineer [Req #5JT1111]** Responsible for SW dvlpmt (C, Python, etc.) of SSD test tools that are req'd for in-house validation & regression testing of SAS and NVMe SSD's. **Staff SW Engineer [Req #5JE1611]** Responsible for Linux device driver dvlpmt of (SAS) & (NVMe) for Storage Intelligence features. **Sr. SW Dvlpmt Engineer [Req #5KL1007]** Conduct research & dvlpmt & novel memory & storage SYS SW research with existing memory technologies & emerging memory technologies. **Research Engineer, Senior [Req #5KL2013]** Research new computer vision 7 machine learning algorithms for advanced product prototypes. **Sr. Algorithm Engineer [Req #5KL2512]** Invent & refine optimal real-time algorithms for physiological signal processing as well as the associated SW infrastructure for building the mathematical models.

Mail resume to Samsung Semiconductor, c/o Staffing – PTCL, 3655 North First Street, San Jose, CA 95134. Must reference Req# to be considered. EOE

Help build the next generation of systems behind Facebook's products.

Facebook, Inc.

currently has the following openings in **Menlo Park, CA (various levels/types)**:

Software Engineer (SWEBNOVJ) Create web &/or mobile applications that reach over one billion people & build high volume servers to support our content. Bachelor's degree required. Exp. may be required depending on level/type. **Software Engineer (SWEMNOVJ)** Create web &/or mobile applications that reach over one billion people & build high-volume servers to support our content, utilizing graduate level knowledge. Master's degree required. Exp. may be required depending on level/type. **Engineering Manager (3334J)** Drive engineering effort, communicate cross-functionality, & be a subject matter expert; &/or perform technical engineering duties & oversee a team of engineers. **Research Scientist (5239J)** Research, design, & develop novel visualization systems to make data actionable & facilitate decision making at Facebook. **Product Designer (6064J)** Design, prototype, & build new features for Facebook's website or mobile applications. **Engineer (6256J)** Implement algorithms in embedded software on target production hardware. **UX Researcher (5668J)** Oversee & design the user experience component to generate actionable insights. **Manager, Partner Engineering (2583J)** Work with our strategic partners to integrate Facebook Platform into their websites, applications, & devices. Domestic &/or international travel required to various unanticipated locations. **Abuse Analyst 5958J)** Review & escalate anomalous computer & system cybersecurity alerts/activity. **Lead, Product (957J)** Plan business objectives, develop product strategies & establish responsibilities across product area. **Product Manager (6182J)** Engage in product design & technical development of new products. **Data Engineer, Analytics (3341J)** Build data warehouse plans for a product or a group of products. **Principal, Auction and Delivery Science (6017J)** Perform custom financial & economic performance analysis & share resulting insights with Product teams & select advertisers.

Mail resume to: Facebook, Inc. Attn: SB-GIM, 1 Hacker Way, Menlo Park, CA 94025.
Must reference job title & job# shown above, when applying.

NEW IN 2015

IEEE TRANSACTIONS ON
BIG DATA

SUBSCRIBE AND SUBMIT

For more information on paper submission, featured articles, call-for-papers, and subscription links visit: www.computer.org/tbd

TBD is financially cosponsored by IEEE Computer Society, IEEE Communications Society, IEEE Computational Intelligence Society, IEEE Sensors Council, IEEE Consumer Electronics Society, IEEE Signal Processing Society, IEEE Systems, Man & Cybernetics Society, IEEE Systems Council, IEEE Vehicular Technology Society

TBD is technically cosponsored by IEEE Control Systems Society, IEEE Photonics Society, IEEE Engineering in Medicine & Biology Society, IEEE Power & Energy Society, and IEEE Biometics Council

IEEE

IEEE computer society

Apple Inc. has the following job opportunities in Cupertino, CA:

Software Engineer Applications (Multiple Positions Open) (REQ# 9FLQZC). Des, dev, and maintain complex front-end code for the Apple Online Store.

Software Development Engineer (REQ#9S3S4P). Dev web-based tools to improve the qual & exp of Apple Maps, incl. client services & data pipeline.

Software Development Engineer (REQ#9UK48T). Bld a platform to fully utilize probe and third party data to improve the Maps exp.

Software Development Engineer (REQ#9TLRM7). Des & implem test automation framework and test cases to test battery life of dev's running iOS, incl iPhone & iPad.

Panel Technology Project Lead (Hardware Dev Eng) (REQ#9BP2FP). Resp for leading overall tech schedule and milestones, driving project execution. Travel Required 30%.

Software Engineer Applications (REQ#9NKTEW). Des, dev & debug SW for hyper-scale servr envmts.

Software Engineer Applications (REQ#9Y8V8K). Design & dev SW for app deployment automation.

Product Design Engineer (REQ# 9HCN7V). Visualize, design & produce new Apple products. Travel req'd: 20%.

Software Engineer (REQ#9WRQY8). Design & validate SW application for customer sytems.

Hardware Development Engineer (REQ#9E62WE) Des & dvlp camera modules for consumer electronic devices. Travel req. 20%.

Software Development Engineer (REQ#9DCRNT) Dsgn & dvlp core Apple frameworks for mobile iOS devices.

ASIC Design Engineer (REQ# 9FETMH) Dsgn verif of complex SOCs for consumer prdcts.

Unix/Linux Systems Engineer (REQ# A2L2EH) Install & coordinate Apple's manufacturing critical systems.

Software Development Engineer (REQ#9A7TMJ) Debug Cantonese-lang specific natural lang & speech recogn techs.

Software Development Engineer (REQ#9FFR35). Design and dev. digital personal assistant for mobile dev.

Engineering Project Specialist (REQ# 9ZK3C5): Plan, coord, ensure resource avail, & track prog's for lrg & complex projs w/i Apple's retail Point of Sale system. Travel Req 20%.

Software Quality Development Engineer (REQ#A2758H). Identify issues, test features, & drive qual of Siri user exp.

Software Engineer Applications (REQ#9W2N5C). Respon for building sw & sys to manage infrastructure & apps through automation.

Software Development Engineer (Multiple Positions) (REQ#9MGW5R) Dsgn, dvlp, & debug SW to control NAND flash devices by implmnting FW for ASIC controllers.

Software Development Engineer (REQ#9BJ2XJ). Tst, brk & imprv Apple's revltnry mobl pymnt soln.

Systems Design Engineer (REQ# 99UW4E). Perf RF paramtrc qualif & reqs adhrnce eval on Apple prods.

Software Development Engineer (REQ#9E9PMN). Responsible for dev SW for analytics data collection & analysis.

Software Engineer Applications (REQ#9F224H). Respons for the des, implem, & maint of robust cross-platform SW solutions in order to meet Apple's content protection needs (server, desktop, & embedded).

Product Design Engineer (REQ# 9FNPUJ) Analyze & pvde des feedback for air flow through con elect dev while using expertl & simul tech.

Software Development Engineer (REQ#9DCTJY) Res for the des and dvlpmt of algorithms & SW for comp vision sys.

Information Systems Engineer (REQ# 9U22A5) Resp for admin, instal, config, trbleshoot & write of supp doc of IBM/AIX HW tech.

Software Engineer Applications (REQ#9NBSCP) Des high-scale web svs for iCloud & related apps w/in the Apple ecosys.

Software Engineer Applications (REQ#9VSUQJ) Des & dev SW for analytics platform.

RF Test Engineer (REQ#9EGPD9). Plan & execute New Prdct Intro to world wide mnfctring sites w/a focus on RF testing. Travel req'd 25%.

Operations Engineer (REQ# 9MC2BL). Collab w/ Suppliers & Eng teams to dev the manufacturing blueprint for Eng builds, such as process flows equip/fixture des, implementation schedules & validation plans.

Software Development Engineer (REQ#9HHSC9). Des and dev SW for real-time road traffic estimation.

Software Development Engineer (REQ#9REU7R). Dsgn & dvlp highly scalable dstrbtd sys to process large scale data using Java, Scala & Map-Reduce frmwrk.

Software Development Engineer (REQ#9M5U4E). Dev & implem user interface & interaction SW on iOS and MacOS platforms for media editing apps

Software Engineer Applications (REQ#9GAVQV). Des, dev, & manage SW & tools for large-scale sys op's & deployment automation.

Software Development Engineer (REQ#9JPV7L) Test the support Bluetooth SW in Apple prod & ens com w/the various Bluetooth accessor rel in the mkt. Travel req. 15%.

Software Development Engineer (REQ#9L8QHR) Monitor, triage, & respond to security issues in prdcts for sale, online services offered, & each of the SW components that make up the operating sys, apps, and srvces.

Product Design Engineer (REQ# 9E62RJ) Work as part of the cross-functional dsgn team responsible for conceiving, desgning, and prdcing exciting new prdcts. Travel req. 20%.

Software Applications Engineer (REQ#9NRTLY). Des & dev highly scalable server SW for the Contacts, Calendars, & Bookmarks serv's w/

complex DB interaction, mntn & enhance existing app's & framewks.

Software Development Engineer (REQ#9UBVA2). Des and dev SW and test cases for testing map SW.

Hardware Development Manager (REQ#9GWQSP) Dsgn & dvlp tightly integrated antenna systms for portable devices, including mobile phones & wearable devices. Travel req. 20%.

Software Engineer Applications (REQ#99XUSY). Rsrch, dsgn, dvlp, implmt & debug machine learning models for iTunes & App store rcmmndtn models to power the prod exprnce.

Re-Engineering Specialist (REQ# 9T7VBU). Use Math/Statist modeling & Ops Research techqs to dev innovative sol'ns, provide modeling & data anlysis expertise, & support strateg & tactical decision making.

Software Development Engineer (REQ#9VL27H). Des & dev SW for WebKit web brwsr tstng & toolng projs.

Hardware Development Engineer (REQ#9H8QKF). Resp for perf acoustic sim at module & sys lvls.

Systems Design Manager (REQ# 9FTNWS). Des, dev, deploy & spprt wirelss Connctvty statns for calibrtn and tstng of Apple prods. Travel req'd: 20%.

Engineering Project Lead (REQ# 9E632Z). Resp for managing & commnctng status & issues of ongoing wrless eng projs.

ASIC Design Engineer (REQ# 9FESYW). Des anlg & mix-sig integ circts for wrless comm.

Strategic Sourcing Manager (REQ# 9GUV3N). Dev strat mem commod plan for new prod intro. Travel Req'd: 20%.

Software Quality Assurance Engineer (REQ#9JFPF2). Conduct SW testing of iTunes Store feature & contents.

Interactive Content Producer (REQ# 9PNSL9). Wrk cross-functionally across mltple teams. Directly participate in the evolution of the content mngmnt systm platform & prfrm user acceptance testing of systms & toolsets.

PC Board Designer (REQ#9X6SDG) Study PC Board (PCB) design trends and processes.

Data Engineer (REQ#9XTPCW). Prgram data extract, anlysis, reportng & automation by establishng & utilizng infrastructure & intel tools to provide data solutions for biz.

Software Development Engineer (REQ#9RMND6). Perform initial design & dev of apps.

Software Development Engineer (REQ#9DSQX3). Dev Big Data analytics apps for CoreMedia, FaceTime, GameCenter & their underlying netwrking archit.

Senior Quality Assurance Engineer (REQ#9VNPYA). Dsgn & dev spprt libraries & infrastructure to spprt SW test'g for ntwrk'g subsystems & spprt oprt'g sys functionality.

Software Engineer Systems (REQ# 9BGW2Y). Idntfy and reslve WLAN-reltd issues across iOS & OS X pltfrms.

ASIC Design Engineer (REQ#9LTVF8) Des/dev complex, high perf, ultra-low-power HW SOC containing many clock & voltage domains for consumer prods.

Software Engineer Applications (REQ#9PBMDT) Des/dev SW solutions in sup of Apple Online Store.

Systems Design Engineer (REQ# 9F4R55). Dev product sustainability & environ. specifications & procedures.

Software Engineer Applications (REQ#9PBMEP). Des & dev SW for Customer Systems group in service mgmnt domain.

Software Engineer Applications (REQ#9MQVAA). Enhnce & maintn iTunes Store new commrce orientd projects of existing commrce feats.

Software Systems Engineer (REQ# 9FKPK6). Des & dev SW to supp connectivity over USB & other transports.

Apple Inc. has the following job opportunities in San Francisco, CA:

Software Development Engineer (REQ#9LUQHE). Des and execute testing strategies for large scale Internet products.

Apple Inc. has the following job opportunities in Austin, TX:

Technical Project Lead (REQ# 9GMSFN). Under gen suprv, dev opor readiness plan for new prod intro.

Apple Inc. has the following job opportunities in Maiden, NC:

Network Engineer (REQ#9F7UB9). Respons for des, implem, & support of Apple's global data center network infrastructure.

Refer to Req#
& mail resume to
Apple Inc., ATTN: L.J.
1 Infinite Loop 104-1GM
Cupertino, CA 95014.
Apple is an EOE/AA m/f/ disability/vets.

TECHNOLOGY

Intuit Inc.

has openings for the following positions in **Santa Clara County**, including
Mountain View, California or any office within normal commuting distance:

Software Engineers (Job code: SW1115): Apply software development practices to design, implement, and support individual software projects. Work on problems of moderate scope and complexity where analysis of situations or data requires a review of multiple factors of the overall product and service. Review product requirements and architecture to understand and implement software projects. **Senior Software Engineers (Job code: SSW1115):** Exercise senior level knowledge in selecting methods and techniques to design, implement, modify and support a variety of software products and services to meet user or system specifications. Work on problems of complex scope where analysis of data requires evaluation of multiple factors of the overall product and service. Analyze and synthesize a variety of inputs to create software and services. **Staff Software Engineers (Job code: STSW1115):** Apply master level software engineering and industry best practices to design, implement, and support software products and services. Evaluate the most relevant factors and exercise independent judgement in the creation, design, implementation or modification of software and services. Act as a technical lead for complex projects. **Software Engineers in Quality (Job code: SWQ1115):** Apply best software engineering practices to ensure quality of products and services by designing and implementing test strategies, test automation, and quality tools and processes. Work within a moderate scope, covering a range of technologies and level of complexity where analysis of situations or data requires a review of multiple factors to ensure quality of the overall product or service. Review product requirements and architecture to create and implement quality engineering requirements. **Senior Software Engineers in Quality (Job code: SSWQ1115):** Apply senior level software engineering practices and procedures to design, influence, and drive quality and testability of products and services. Work within complex scope, covering a range of technologies and level of complexity where analysis of situations or data requires a review of multiple factors to ensure quality of the overall product or service. Exercise judgment in application of methods and procedures to ensure quality products and services. **Senior Technical Data Analysts (Job code: I-327):** Leverage technical skills, business acumen, and creativity to extract and analyze massive data sets, build analytics-ready datasets to surface insights and key business metrics. Sift through large amounts of information, extract key insights and communicate succinctly. **Senior Technical Data Analysts (Job code: I-384):** Lead initiatives to collect, interpret, and report on key business metrics. Apply skills and systems expertise to create reports and analysis that provide actionable insights to business stakeholders. **Senior Business Analysts (Job code: I-1778):** Partner with marketing, finance, analytics, and cross-functional teams to interpret large volumes of data, address key business questions, from hypothesis to execution.

Positions located in **Plano, Texas**:

Software Engineers (Job code: I-151): Apply software development practices to design, implement, and support individual software projects. Work on problems of moderate scope and complexity where analysis of situations or data requires a review of multiple factors of the overall product and service. **Senior Software Engineers in Quality (Job code: I-191):** Apply senior level software engineering practices and procedures to design, influence, and drive quality and testability of products and services. **Senior Strategy Managers (Job code: I-2045):** Engage with key stakeholders across multiple projects and work streams. Develop an understanding and stay current with the competitive landscape.

To apply, submit resume to Intuit Inc., Attn: Olivia Sawyer, J203-6, 2800 E. Commerce Center Place, Tucson, AZ 85706.
You must include the job code on your resume/cover letter. Intuit supports workforce diversity.

TECHNOLOGY

LinkedIn Corp.

LinkedIn Corp. has openings in our **Mtn View, CA** location for:

Software Engineer (All Levels/Types) (SWE1115MV) Design, develop & integrate cutting-edge software technologies; **Test Engineer (6597.1067)** Drive test automation of products, working within a highly distributed platform capable of scaling beyond 380 million members, serving billions of page views; **Associate Product Manager (6597.737)** Develop strategy for a product areas & features at LinkedIn; **User Experience Designer (6597.876)** Create innovative end-to-end user experiences for existing & new product offerings; **Manager, Software Engineering (6597.45)** Lead a group of Software Engineers responsible for building & supporting a diverse set of tools used by Developers to build & test products; **Data Warehouse Engineer (6597.1227)** Design, build, & maintain all parts of the data warehouse infrastructure; **Test Engineer (6597.1167)** Design & develop advanced test suites & necessary automation frameworks using object-oriented methodologies.

LinkedIn Corp. has openings in our **Sunnyvale, CA** location for:

Software Engineer (All Levels/Types) (SWE1115SV) Design, develop & integrate cutting-edge software technologies; **Staff Network Engineer (6597.113)** Responsible for the design & implementation of highly reliable, scalable & secure, high performance network solutions; **Talent Integration Solution Developer (6597.1460)** Design, develop, test, deploy, support, & enhance custom web portals & integration solutions & seamlessly connect LinkedIn enterprise systems; **Senior Test Engineer (6597.1030)** Write & build automated test suites, while continuously designing creative ways to break software & identify potential bugs.

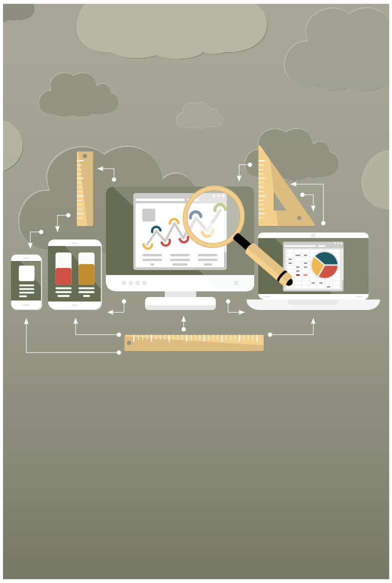
LinkedIn Corp. has openings in our **San Francisco, CA** location for:

Software Engineer (All Levels/Types) (SWE1115SF) Design, develop & integrate cutting-edge software technologies.

LinkedIn Corp. has openings in our **New York, NY** location for:

Software Engineer (All Levels/Types) (SWE1115NY) Design, develop & integrate cutting-edge software technologies.

Please email resume to: 6597@linkedin.com. Must ref. job code above when applying.



Executive Roundtable Series

Machine Learning and Cognitive Computing

Seth Earley, *Earley Information Science*

Machine learning and cognitive computing are today's newest buzzwords, and there is a lot of hype surrounding them in the market. This article is based on a recent webinar on analytics produced by *IT Professional*, the *Journal of Applied Marketing Analytics*, and consultancy Technology Business Research (TBR), along with the Content Wrangler; it was hosted by Earley Information Science (EIS). Video of the webinar is available at https://youtu.be/8Tx4GuSU_Wc. The goal is to help organizations understand what's practical and what's possible in the fast-growing fields of machine learning and cognitive computing, and how these fields are related to artificial intelligence (AI).

Earley: Let's first talk a little bit about machine learning, cognitive computing, the AI that underlies them, and where those fields have been going. Machine learning is a technique for detecting patterns and surfacing information, using many different mechanisms based on statistics and mathematical models. One good example is search technology, which provides entity extraction, clustering, and classification. Machine learning algorithms pull back information that's relevant to the user by looking for patterns, improving the search results. For example, Amazon uses machine learning algorithms to provide suggestions about other products that are related to the ones a user has viewed or selected.

Cognitive computing is a newer, emerging field. It's about making computers more user friendly, with an interface that understands more of what the user wants. It takes signals about what the user is trying to do and provides an appropriate response. Siri, for example, can answer questions but also understands context—whether the user is in a car or at home, moving quickly and therefore driving, or more slowly while walking. This information contextualizes the potential range of responses, which are therefore more personalized.

AI encompasses all of these tools and solves a wide variety of problems—everything from writing articles to driving cars, detecting fraud, and diagnosing diseases. A lot of decisions that have typically been made by human beings can now be made by means of AI. So, there are many developments in the AI field that are practical and actionable.

With this background in mind, what are your thoughts on this technology?

Daley: The fundamental concepts of AI have been around for 60 years, including the idea of modeling machines to mimic human intelligence. What's really changed in the last few years is the improvements using probability and statistics. The algorithms for deep learning, in particular—which is really a form of backward chaining neural nets [an approach that uses learning algorithms to progressively infer a pattern about a body of data by starting with the goal and then determining the rules that are inferred, which can then be used to reach other goals]—has advanced significantly. Second is the recent discovery that using graphical processing unit [GPU] chips to do calculations in paral-

 Published by the IEEE Computer Society 2376-113X/15/$31.00 © 2015 IEEE

lel allows results that would have taken weeks to be achieved to now be obtained in just a few hours. Finally, and I think probably most significantly, is the amount of data that's available to train AI systems. They need a lot of data, and it's available digitally in all forms. In Google's famous cat experiment, an AI system was exposed to millions of thumbnails from YouTube, and without being supervised, the system was able to recognize a cat.

Downs: Many machine learning solutions have already been developed, and they are continually being improved. I spent some time at Microsoft Research doing some early work in Bayesian reasoning and machine learning. We built a solution for traffic modeling that was spun out as Microsoft Research's first startup company, called INRIX, which now provides real-time and predicted traffic information around the world.

I see three tiers of commercial engagement with these types of technologies. For one group of companies, such as Google, Amazon, Facebook, Microsoft, and Apple, these technologies are strategic, and their investment is a hundredfold or more than it would be from a more conventional business. Second, for some companies, the impact of these technologies could be strategic, but they cannot make the investment in terms of human resources and tools. Finally, we come across companies that have had some of the tools become more accessible to them, but they lack the know-how to use them. Some of them say, "We've tried machine learning, and it doesn't work."

One of the key propositions for machine learning and cognitive computing to really cross the chasm in terms of driving business value is to be able to package up some of these capabilities, perhaps

Roundtable Participants

Bruce Daley is a principal analyst contributing to Tractica's Automation & Robotics practice (www.tractica.com). He focuses on artificial intelligence and machine learning for enterprise applications. Daley has extensive experience as an industry analyst, writer, and publisher focused on the global IT market; has been widely quoted as an industry expert in major publications, including the *Wall Street Journal*, the *New York Times*, the *Financial Times*, the *International Herald Tribune*, *IEEE Spectrum*, and the *San Jose Mercury News*; and is the author of a soon-to-be-published book on data storage, *Where Knowledge is Power, Data is Wealth*. He received a BA from Tufts University. Contact him at bruce.daley@tractica.com.

Olly Downs is chief scientist at Globys. He is responsible for the analytics strategy, technical approach, and algorithm design and development for Globys's marketing personalization technology platform (Amplero). Downs is a machine learning scientist and serial technology entrepreneur, credited with bringing advanced analytics and machine learning methods to bear as the creative spark behind numerous early stage technology companies. He has a PhD in applied and computational mathematics from Princeton University. Contact him at odowns@globys.com.

Mitchell Shuster is an informationist and data scientist at Knowledgent Group (www.knowledgent.com). He specializes in applying advanced analytics and data science concepts and techniques, including machine learning (regression, neural nets, support vector machines, clustering, PCA, anomaly detection, and so on), to help client organizations gain actionable insights and competitive advantage. Shuster previously designed and developed the basis for Intel's worldwide high-volume manufacturing at the newest technology node and was recognized for computational modeling and process implementation. He received a PhD in physics and multiple research fellowships from Penn State University, where he authored research published in multiple prominent peer-reviewed scientific journals. Contact him at mitchell.shuster@knowledgent.com.

Patrick Heffernan is practice manager and principal analyst in the Professional Services Practice for Technology Business Research (TBR; www.trbi.com). He covers the areas of IT services, management consulting, global delivery, strategy and operations, cloud, intelligence cycle, project management, and client engagement. Heffernan received an MA in foreign affairs from the University of Virginia. Contact him at patrick.heffernan@tbri.com.

Seth Earley (moderator) is CEO of Earley Information Science (www.earley.com). He's an expert in knowledge processes, enterprise data architecture, and customer experience management strategies. His interests include customer experience analytics, knowledge management, structured and unstructured data systems and strategy, and machine learning. Contact him at seth@earley.com.

with some specific applications in mind, that actually are able to add value without requiring hands-on intervention. The tools can act without human expertise and provide information on what the technology is discovering.

Earley: How is this technology being applied in marketing?

Downs: Without technological solutions, marketing is a relatively slow process that engages at the enterprise level and uses a large number of resources. It isn't unusual to have 50 or a hundred people involved in planning, campaign formulation, and testing. At Globys, we allow the marketer to drive the discovery and continual optimization both from a large and price value sensitivity and to compare components automatically. The machine learning actually discovers and reads out back to the marketers what the right target audience is for an offer or new product. It makes the process more efficient and less labor-intensive.

Earley: Of course, there is no magic bullet; we need to start with offers and a hypothesis of some sort, as well as have clean data inputs.

Downs: Yes, that's correct. Certain foundational elements need to be in place.

Earley: What cautions would you propose for companies in using machine learning?

Schuster: Machine learning is a tremendously powerful tool for extracting information from data. But as Seth said, it's not a magic wand. You can't simply say "machine learning" and make your problems go away. You have to frame the questions in a way that really allows the algorithms to answer them. Data needs to be set up in the appropriate way, and that can be complex. Sometimes, the data needed to answer the questions may not be available, which can be another barrier. Once the results come out, they need to be interpreted. It's very important to understand the context. The marketing algorithm can tell a marketer what's working the best, but the marketer still needs to know how to leverage the information.

Earley: What concerns do companies have about providing machine learning services?

Heffernan: I see several levels of concern in the machine learning field; some are from the vendors, some are from the prospective customers, and some are from the IT staff. The IBMs and Accentures of the world are afraid of getting behind their competitors. As much as they understand the potential, they aren't sure they can make a compelling business case or that they have the resources to provide the services. For the IT services vendors who truly rely on IT for their business success, having the right people makes all the difference. Without the right staff, they are also-rans.

The second type of concern comes from potential clients, the businesses that could benefit from machine learning and cognitive computing. The real fear here is pretty simple. Can I afford it, or can I afford not to do this? Am I already competitively behind? Am I throwing good money at a pointless project? I think when you take a step back, everyone truly appreciates the potential of groundbreaking technology. Every C-suite decision maker has an iPhone and understands the truly radical nature of IT in our personal and business lives now, compared to 15 years ago. You don't need to make the case anymore about how important technology is and what it can do. The challenge still remains, though—is this the right investment?

The third fear is from employees who are wondering what happens to their work when the machines take over. We've done some research into robotic process automation in business process outsourcing, which sounds a lot like word salad; but what it basically means is the robots doing the work in a call center, for example. We've seen that one robot can replace three people and have an immediate impact on cost, a reduction of as much as 25 percent. It's not just call center staff, but IT staff as well. As we've seen in digital and cloud and business intelligence, it's the lower-skilled IT services folks who are in the most trouble.

Earley: What about the higher-end skills? How does that issue fit into the machine learning picture?

Heffernan: At that end of the skills scale, you have the data scientist with an MBA and seven years of industry experience—what we sometimes call a "people unicorn." That's the person who can translate business needs into machine learning possibilities, then to creating the analytics, and from there, back to the business. That person is in huge demand right now. Being able to find that top talent, retain it, and manage it is really what makes all the difference. One company that's embedding cognitive computing into its business process outsourcing offerings is in fact laying off people at the lower levels, or they're being retrained for jobs that require higher-level services.

The employees who are staying with the cognitive computing-enabled BPO offers are the ones who can formulate a strategy to get tested, translate the business problems into analytics, and communicate the results. We're looking at an increased premium on IT staff at the more advanced levels. The premium was already high, and it's only getting higher. These companies have to make a choice. Are they going to invest in cognitive computing? IBM's Watson is one example of a major commitment. Wipro has made cognitive computing an important pillar going forward, and they are truly building out their own capabilities. Other companies are partnering to provide those high-skilled, high-priced, high-margin, high-value consultants and leaving the more basic IT aspect of it to others. We're looking at this space as one that might completely upend existing business models and how companies are run.

Earley: So how do organizations get started, if their business models are going to change so dramatically?

Daley: The first step is almost always to get the people who have business problems to be solved into a room to discuss their needs. In the past, typical line-of-business leaders didn't really care what the IT solution was—they just wanted the problem solved. I think that's changed now, and a lot of business leaders are much more familiar with the technological possibilities. What they need is guidance, with understanding of how to apply their knowledge to their business problems, and to understand what is practical to actually get done.

One possible strategy is to start working with an open source AI package to understand the technology. With "black box" AI solutions, you may not know why someone got turned down for credit. It's just, "The machine turned it down." If the IT department gets some experience with open source tools, they can get that transparency and be able to understand a little bit about how AI works, and address the issues of how the software works in the future. There is an opportunity out there, and at least at this juncture, open source seems to be the way to go.

Earley: How do you tell what's practical and achievable for a given organization?

Shuster: A large array of problems can be addressed by machine learning. I wouldn't even say there are any issues that can't be helped by this technology. It comes down to what kinds of data you have and what kinds of questions you want to ask. Are the data sources internal or external, and how much do you trust them? How clean and well organized is the data, and what kind of governance is around it? Then, when you ask the question, you can trust the answer, which I think is a bigger potential barrier than the type of question that can be asked.

Earley: So the table stakes are really a core architecture and clean data to work with. Some would argue that their tools work with messy data; however, I would not think many work well with bad data. Most AI experts will say that no matter the type of system, it will perform better when it's given the products, services, processes, and unique terminology of the business. A domain model is an important part of the puzzle. What kind of culture change is needed to make this transition?

Downs: There needs to be a shift to decide that the business is going to be oriented around analytics, and a kind of "no excuses" approach to understanding at least some of the output. In terms of getting going and understanding what's feasible, I think a lot of the first stab efforts today get bogged down in the infrastructure and other challenges with getting the data in place. They may focus too much on getting IT efficiencies in data warehousing rather than trying to solve a relatively simple but soup-to-nuts problem that can show ROI at the very highest level, where it impacts business functions. That is a better approach than just having a technology activity that impacts technology stakeholders. It's always a challenge, because you have to have a cross-functional initiative to pull that off. But it makes the business value much more apparent to the high-level stakeholders. It also forces the company to become more analytically minded across the business, because you engage multiple functions. You are more likely to get buy-in because you are actually driving some appreciated high-level business value.

Earley: Right. We can't walk into the business and talk about data and architectures. We need to speak to the benefits and how we support specific objectives and solve problems. How do organizations get the education they need to tackle these analytics issues?

Daley: To be honest, I haven't seen a lot of unicorns in my career who can span multiple disciplines. AI is really based on a strong mathematical foundation. I think that's absolutely indispensable. But I'm not sure there is any specific roadmap. Certainly one of the things I've been surprised about in my research about AI is how many papers have been written on the topic—over 150,000 research papers on neural nets over the last 60 years. As a result, what I see is that applications are springing up all over the world, independently of one another. Some will be very successful, and some will not be successful at all. Organizations need to be willing to accept failure as a part of learning, and start experimenting.

Earley: So learning by doing can provide practical lessons; however, the organization has to have some tolerance for exploration. In our experience, the exploration can be done in the context of the problem being addressed, which grounds the exploration in practical outcomes. Otherwise, it's a science project. That's okay if there's budget for pure experimentation; however, there better be some clear application eventually, or that certainly won't last. What other applications have you seen in which organizations have tried projects that provided good learning experiences?

Downs: The US Postal Service actually has one of the oldest deep learning implementations that exists, which is a dedicated piece of hardware for handwritten zip code recognition. They deployed the system in 1987. Of course now, deep learning is a very topical technology, with a lot of focus on image recognition, and researchers are branching out into domains such as medicine, with medical imaging as a developing area.

The Executive Roundtable Series

The Executive Roundtable Series took place over four weeks from 27 May to 17 June. In addition to the Machine Learning and Cognitive Computing webinar presented here, we covered three other topics:

- Mining Business Insights with Big Data Analytics and the Internet of Things
- Metrics for Measuring the Customer Experience and Digital Marketing Success
- Using Business Analytics to Drive Higher ROI and Organizational Change

Look for a summary of these webinars, along with full video, in a future issue of *IT Professional*.

There are many applications in language, both spoken and written, with applications in areas that have a lot of inefficiency, such as medical billing, re-admittance, bill prediction, and fraud detection. Some of the early neural network applications were in finance. The advertising auction marketplace is another good example. An ad is actually bid on in a second-price auction in real time, while the page renders it in your browser within a time span of 300 milliseconds.

Some additional work has been sparked by European privacy laws. This has produced a wave of innovation about how to understand the qualities of individuals and being cognizant of them without knowing their specific identity. A market economy has developed around advertising of all types, whether it is in games, in apps on a mobile device, or online. A better understanding of the customer drives efficiency in the marketplace around advertising and display. It is highly focused, particularly as we push more and more toward anonymized information, which requires increasingly sophisticated analytics.

Earley: This is a fascinating space. The ROI is clear, and it's easy to see the impact of analytics for businesses today and how developments on the horizon will continue to disrupt. Digital marketing is a tremendous application area. In terms of what needs to be in place to make this technology work, you still need an architecture. These technologies are not magic. You do need to understand the data structure, what the rules are, and have the right components assembled for digital marketing to work. You need to understand your customer attributes and have a mechanism for pulling them from multiple sources. The tools do provide unsupervised learning approaches that show patterns, but you will still need to know what you are looking for and what you are trying to get. This aspect of machine learning, along with many other potential applications, provides a tremendous opportunity both over the near term and farther down the road. IT

This article originally appeared in IT Professional, *vol.17, no. 4, 2015.*

Move Your Career Forward

IEEE Computer Society Membership

Explore Related Resources on these 2015 Tech Trends

Create Connections

IEEE Cloud Computing Initiative

To help accelerate the development and use of cloud computing technologies, this initiative offers a comprehensive set of resources and activities that leverage expertise from the across the globe. The first-of-its-kind collaborative effort includes six interdependent tracks: a Web portal, conferences, continuing education courses, publications, standards development, and a testbed.

Advance Your Career

Security Certificate of Achievement

The Computer Society now offers a Certificate of Achievement in the growing field of security. Take advantage of this professional development opportunity to become an expert in your office and advance your career by successfully completing these four courses:

- Foundations of Software Security
- Secure Software Design
- Managing Secure Software Development
- Secure Software Coding

Security Certificate of Achievement

The 12 certificate areas assess a candidate's proficiency in understanding the concepts involved, as well as their ability to apply them in the software engineering process.

FOR DIRECT LINKS TO THESE RESOURCES, VISIT

www.computer.org/edge-resources

IEEE computer society

The Community for Technology Leaders

IEEE COMPUTER SOCIETY: Be at the Center of It All

IEEE Computer Society membership puts you at the heart of the technology profession—and helps you grow with it.

Here are 10 reasons why you need to belong.

1 **Training** that sharpens your edge in Cisco, IT security, MS Enterprise, Oracle and more.

2 **Certifications** and exam preparation that set you apart.

3 **Industry intelligence,** including *Computer,* myCS, Computing Now, and myComputer.

4 **Customized solutions** in software and systems, information and communications technology, security and privacy, and computer engineering.

5 300-plus chapters and more than 25 technical committees **keep you connected.**

6 **Access** to hundreds of books, 13 technical magazines and 20 research journals.

7 **Deep discounts** on magazines, journals, conferences, symposia and workshops.

8 Opportunities to **get involved** through speaking, publishing and volunteering opportunities.

9 **Scholarships awarded** to computer science and engineering student members each year.

10 **A robust jobs board,** plus videos, articles and presentations to help you land that next opportunity.

IEEE Computer Society—keeping you ahead of the game. Get involved today.

www.computer.org/membership

IEEE computer society