

**CIO**

**THINK TANK**

FROM IDG

CIO Think Tank Roadmap Report:

# DATA and ANALYTICS at SCALE

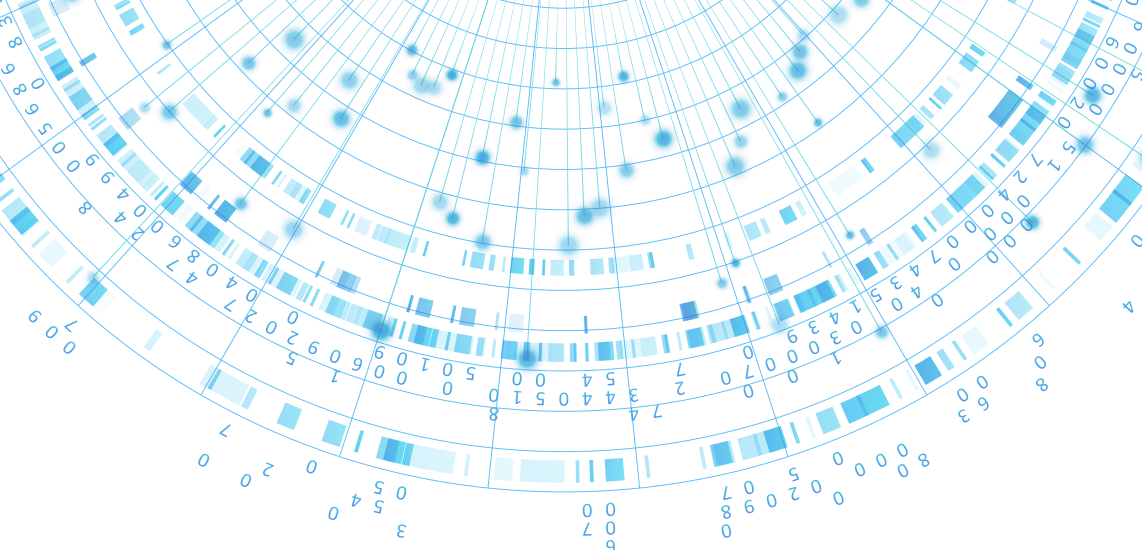
31 IT leaders share challenges,  
strategies, and insights



Sponsored by



**Hewlett Packard  
Enterprise**



# The world is awash in data. So, in almost every enterprise, the question looms: How can we scale data analytics to deliver the greatest possible value?

**In July and August 2021**, CIO held three virtual CIO Think Tank discussions that brought together 31 IT leaders to unpack one of the most important issues in enterprise technology today: maximizing the utility of data collected through multiple channels. The roundtables also included Ritu Jyoti, group VP, Worldwide AI and Automation Research Practice, IDC; Eric Knorr, editor-in-chief, IDG; and data strategy executives Robert Christiansen, Dr. Eng Lim Goh, and Matt Maccaux, HPE.

The goal of these discussions was to identify key challenges in analytics initiatives and to offer a roadmap for IT leadership—as well as the technology industry—to overcome those obstacles. All participants drew on their experience and knowledge to describe strategic and tactical approaches to scaling analytics.

Data now flows through and among all aspects of an enterprise, as well as through customers and partners. This omnipresent reality means that data analytics transcends individual applications, with interwoven ecosystems of people, processes, and technologies. Our CIO Think Tank panelists shared their views about confronting this broad scope of work, including culture, hiring, training, organizational models, governance, and more.

The participating IT leaders hail from a variety of industries (see the Participants list on page 11), each at a different stage of their journey toward democratized, data-driven decision-making. This report consolidates their thinking on key topics to help guide other CIOs and business executives seeking to deliver on the promise of data and analytics at scale.

## TABLE OF CONTENTS

- 2**  
Introduction
- 3**  
CIOs Say Analytics Is a Journey
- 5**  
Challenge 1:  
Changing the Culture
- 6**  
Challenge 2:  
Balancing Strategy With  
Immediate ROI
- 7**  
Challenge 3:  
Building a Single Source  
of Truth
- 9**  
Challenge 4:  
Finding Talent and Skills
- 10**  
Challenge 5:  
Democratizing Analytics
- 11**  
Participants and Thanks



### WHAT IS CIO THINK TANK?

CIO Think Tank is a unique collaboration showcasing the ideas and expertise of top IT executives, IDC analysts, IDG editors, and our exclusive vendor partner. Our goal is to explore and shape the future of the IT function and emerging technologies. HPE is IDG's partner on this Data and Analytics at Scale: CIO Think Tank. For more on this program, go to [www.idg.com/product/cio-virtual-think-tank](http://www.idg.com/product/cio-virtual-think-tank).

# The Journey to Analytics at Scale

**For most CIOs**, analytics projects top the to-do list. The ideal outcome is an array of insights that results in faster, more efficient operations.

For example, Mastercard receives thousands of requests from partner banks needing logo approval to issue new credit cards. Gurpreet Atwal, senior VP at Mastercard, said this approval process was previously slow and manual. Today, an algorithm crunches applicant data to score most requests, kicking exceptions to manual review.

"What used to take us weeks usually takes us minutes now," he said.

Inpro CIO Steve Baumgartner related similar achievements in assessing 10,000 to 15,000 potential bids each month as his company pursues construction projects.

In other cases, companies are looking at more strategic results, potentially affecting entire business models. Many organizations also have an eye out for direct monetization of data itself, now or in the future.

"My interest at this point is to come up with software products that could potentially not provide us just immediate revenue, but provide a way to collect that strategic data, which could be very valuable going into the future," said Subbu Ramanathan, divisional CIO at Brady, a manufacturing company. As examples, he cited Google and Facebook and their monetization of data gained through products that are often given away.

It's an alluring prospect, one that no longer seems within reach of hyperscale IT vendors alone.

"More and more now, people are thinking that data should not be treated as cost anymore, but instead as an investment in a new asset class, eventually with value on our balance sheet," said Dr. Eng Lim Goh, SVP and CTO, High-Performance Computing and AI at HPE.

IT leaders also know that getting there isn't easy, however. "Analytics at scale" is a large, complex hairball of people, process, and technology issues.

Ritu Jyoti, group VP, Worldwide AI and Automation Research Practice for IDC, divides analytics into four layers. The model itself is simple, but each element has many moving parts:

- **DATA:** What you have and how you store it—operational and analytical databases, data lakes, data streams, and more.
- **DATA ENGINEERING:** How you ingest and transform data for consistency and value.
- **GOVERNANCE:** The processes and principles for managing data. Governance is critical for ensuring data integrity and security across the enterprise.
- **INTELLIGENCE:** Data catalogs and resources available to the end user for new and emerging workloads, including AI and machine learning.

Our panelists cited questions and hurdles at each of these layers from data quality, integration, and governance through organizational structure and support. Balancing strategy and tactics in the long and short term is an acute challenge. Hiring, training, and/or building the right skills in a highly competitive market is another difficulty, for both specialists and the broader workforce.



# 64

**ZETTABYTES**  
(64 TRILLION GIGABYTES)

**Estimated total of data generated in 2020** Source: IDC

## The Journey to Analytics at Scale *continued*

“More and more now, people are thinking that data should not be treated as cost anymore, but instead as an investment in a new asset class, eventually with value on our balance sheet.”



— DR. ENG LIM GOH, SVP AND CTO, HIGH-PERFORMANCE COMPUTING AND AI, HPE

All these challenges are interwoven. For example, working toward a "single source of truth"—a common goal expressed across all three CIO Think Tank panel discussions— requires data and human integration, standardized nomenclature, and often, common data classifications across business units. It's a cultural *and* skills challenge.

Analytics at scale is a different beast. HPE's Goh gave the example of one company's attempt to use video analytics for quality control on the manufacturing line.

"The pilot with 100 cameras worked in a warehouse, but when it came to production for 100,000 cameras, the thinking all had to be changed," Goh said.

Panelists also noted that rolling data together from multiple departments or lines of business often runs afoul of inconsistent definitions and classifications, not to mention the tangle of new and legacy systems.

Success is not guaranteed.

"We all have heard horror stories about big business intelligence and data warehousing projects going on and on and on without any ROI," said Jyoti.

Clearly, one strategy does not fit all for achieving analytics at scale. An approach that succeeds at one organization may run into a different set of roadblocks elsewhere.

Our panelists' companies are very much in the middle of transformational work. The following five challenges are specific to analytics at scale across industries and reflect how panelists are working to overcome them.



2%

Estimated percentage of new 2020 data saved or retained into 2021 Source: IDC

# Analytics Challenge 1: Changing the Culture

**Cultural change is always difficult.** But that is exactly what analytics at scale demands.

"My organization is 100 years old. We've been doing some big data analytics and giving them some visual information for the last five years, and we use about 3% of it," said Rick Spugnardi, CIO of Replacement Parts.

"It's the day-to-day things that we do, the organizational processes, that's the hardest part: cultural change," said Jyoti.

"Data-driven culture" sounds simple, but our panelists agreed that it requires much more than basic literacy with data or statistics. Achieving a real data culture also means overcoming territorialism by data owners and building a culture of sharing, a common vocabulary and understanding of the data, and, crucially, consistency in prioritizing data over gut judgement.

"We have now prioritized data literacy for the entire [workforce], because if they don't increase their data literacy and awareness of how you can work with data in modern ways with more embedded and predictive analytics, then they're still getting a PowerPoint with some BI report that they ran, and nothing happens at the end of it," said Shobie Ramakrishnan, chief digital and technology officer, Pharma Commercial, GlaxoSmithKline.

Several IT leaders also mentioned that analytics at scale requires a culture of agility and speed—as Ramakrishnan put it, "taking on the principles of agile and product management that is so prevalent in the software space and driving it into how we explore data solutions."

“One of the more serious obstacles is not about technology. It's about getting humans to talk to each other about the meaning of the data.”



— ROBERT CHRISTIANSEN  
VP STRATEGY, HPE



## RECOMMENDATIONS:

For organizations that lack top-down support for data analytics, success starts with showing business value, mixed with "the whole risk management thing: Here's the [compliance] impact that could happen and the probability" if the house is not in order, said Tonya Digiorno, IT director at El Dorado County in California.

More commonly, IT leaders need to build the demand and processes for data at a grassroots level in lines of business.

To practically implement a top-down data-sharing mandate at the State of Michigan, Paul Groll, director of Emerging Technology Research, said the state "focused on culture and assigned a chief data officer for every agency, and then chief data stewards." With 20 agencies, Groll said the state has "somewhere north of 400 separate lines of business, so there's a lot of pockets of data here and there."

To further foster collaboration, they created a standard "data sharing contract" to help any two agencies fully understand roles, responsibilities, and limits. The first contract took months as the participants ironed out details, but Groll said the contract has proven an efficient tool.

Building a data culture is an ongoing task.

"We did our first inaugural data conference, where our CEO opened the conference with me and said how much he values data in the company and why that's going to be a differentiator," said Ramakrishnan.

Several panelists mentioned the need for "cheerleading," even at companies that have an established analytics practice.

# Analytics Challenge 2: **Balancing Strategy With Immediate ROI**

**At scale**, governance is paramount to ensuring data integrity and consistency, regulatory compliance, and much more. Executives at various companies emphasized that they can't sit still while formalizing a comprehensive plan; projects and use cases must keep moving ahead.

"As a startup, our motto is to be scrappy, but in the data and analytics world, sometimes you need to take a balance between what is scrappy and what is moving too fast. So that's a challenge, the balancing act," said Alex Mendoza, director, Data and Analytics at Chobani LLC.

Moving too fast can build technical debt and integration challenges. Moving too slowly risks falling behind competitors and failing to deliver business value.

GlaxoSmithKline's Ramakrishnan underscored the point, noting that she has not created a formal, centralized data strategy.

"We've set up a federated operating model, so each function can move at pace, but we have the 20 or 30% that we think are common enterprise priorities—like a data platform, for example," she said.

One reason for this approach, she said, is that the company's operating units (research, manufacturing, and sales, for example) don't share a single "customer experience" that might benefit from more centralized governance.

She said the company may roll out more centralized planning as time goes by, but "we've pivoted to creating value first and then letting some of these things follow. Because now the story becomes much more powerful. We're able to tell real-life stories with real-life examples."



# \$9.01

**Average return per dollar invested in analytics projects**

Source: Nucleus Research, 2019



## RECOMMENDATIONS:

For companies early in the analytics journey, panelists agreed that the imperative is to get started, with a focus on finding early wins. One way to do that is to ask new business questions about the data you already have.

"Here are things that we would like to do to improve the efficiency of our manufacturing, of our supply chain. Here is the data that we already have. What can that tell us by applying new models?" said Brady's Ramanathan, describing the company's starting point. "Then, in addition to that, we're also starting to identify gaps where we don't have data, and we're trying to fill that gap," he said.

For companies still in the earliest stages of scaling their analytics, Jyoti recommended starting with tracking and measuring results, and then communicating them broadly. Data visualization can also help raise enthusiasm and get conversations rolling about key goals and guiding principles.

Finding existing common ground across multiple units or departments provides a basis for faster wins, while potentially mitigating technical debt from new systems or integrations.

In the absence of a full-blown strategy, communicating simple objectives or clarifying questions can rally stakeholders to work in a mutually beneficial way. This sets the stage to develop more formal guidelines later.

"If a citizen comes up and asks me 'What data do you have on me?' I want to be able to answer that question to make sure that people understand how their data is being used," said Digiorno, who was previously CIO for the California Department of Technology, a much larger organization. "So my goal is to come and show that it can be done at a smaller scale, and then you can scale it out at a larger scale once you're successful."

# Analytics Challenge 3: **Building a Single Source of Truth**

**Without consistent data that's consistently interpreted,** business challenges loom.

Rebecca Meyer, IS director of Commerce Apps and Ecommerce for Kelly Moore Paint, described her company's journey to integrate and clarify its data sources.

"The numbers are still going to produce slightly different reports based on how we view and categorize our customers and our territories," she said. "So, I can see that sales team members are always going to have those question marks around. 'Is this the true data, or is there something better that shows I made my numbers?'"

This suggests that, although many companies express a desire to move to a single source of truth, that's often not an absolute. At scale, there's often a balancing act between localized needs and a host of other factors. With the ongoing shift to multicloud operations and SaaS applications, providers hold an increasing amount of data, said Pavi Agrawal, chief applications architect at Mars.

"We've got Azure and Oracle, and we also have Snowflake to contextualize the data and make it business-ready," said Michael Crowley, VP of Technology and Innovation at Martin-Brower. Furthermore, he said, the company's field operations produce a lot of "IoT-type data" that isn't yet ingested into the data warehouse.

Tim Dion, chief data officer for Armor Healthcare, faces similar obstacles as his company delivers health services to prison inmates.

"In the correctional institution world, connectivity [among data sources] is sometimes hard. So, medical records are behind locked doors, and you've got a lot of the data that by law is only resident inside of a specific facility," he said.

Companies are also wrestling with new unstructured data types, which, according to Jyoti, tend to be very disparate and very distributed.

"Last year, we all saw that explosion of more conversational data; more leveraging of unstructured content through decision-making is becoming front and center," she said.

This is another area requiring new knowledge and skills.

"IT generally is all about structured data. But unstructured data ... that capability is clearly on the business side, and now we are working to patch it all together," said Raja Gangavarapu, VP, Wolters Kluwer.

The technical, legal, and definitional issues facing data integration are substantial.

**“**You can't implement good solutions if you don't understand the data.**”**



— TONYA DIGIORNO  
IT DIRECTOR  
EL DORADO COUNTY

# Building a Single Source of Truth *continued*



## RECOMMENDATIONS:

Several IT leaders at large enterprises described a "federated" data model.

"We have enterprise data, and then we have a localized version of truth," said Agrawal.

This approach allows for central financial reporting, while each business segment also has embedded analytical capabilities and retains the data deemed essential.

Mastercard's Atwal cited different reasons for a similar mindset: "Privacy regulations are complicating getting all of this data into a single place. So, we need to design our analytic practices for a decentralized [environment] where all data is not in one place, while still having the need to bring the assets together to get better insights and a competitive advantage in the marketplace."

In terms of building the technical connections, hoping for vendors to solve the problem might be tempting. One panelist asserted that will take too long.

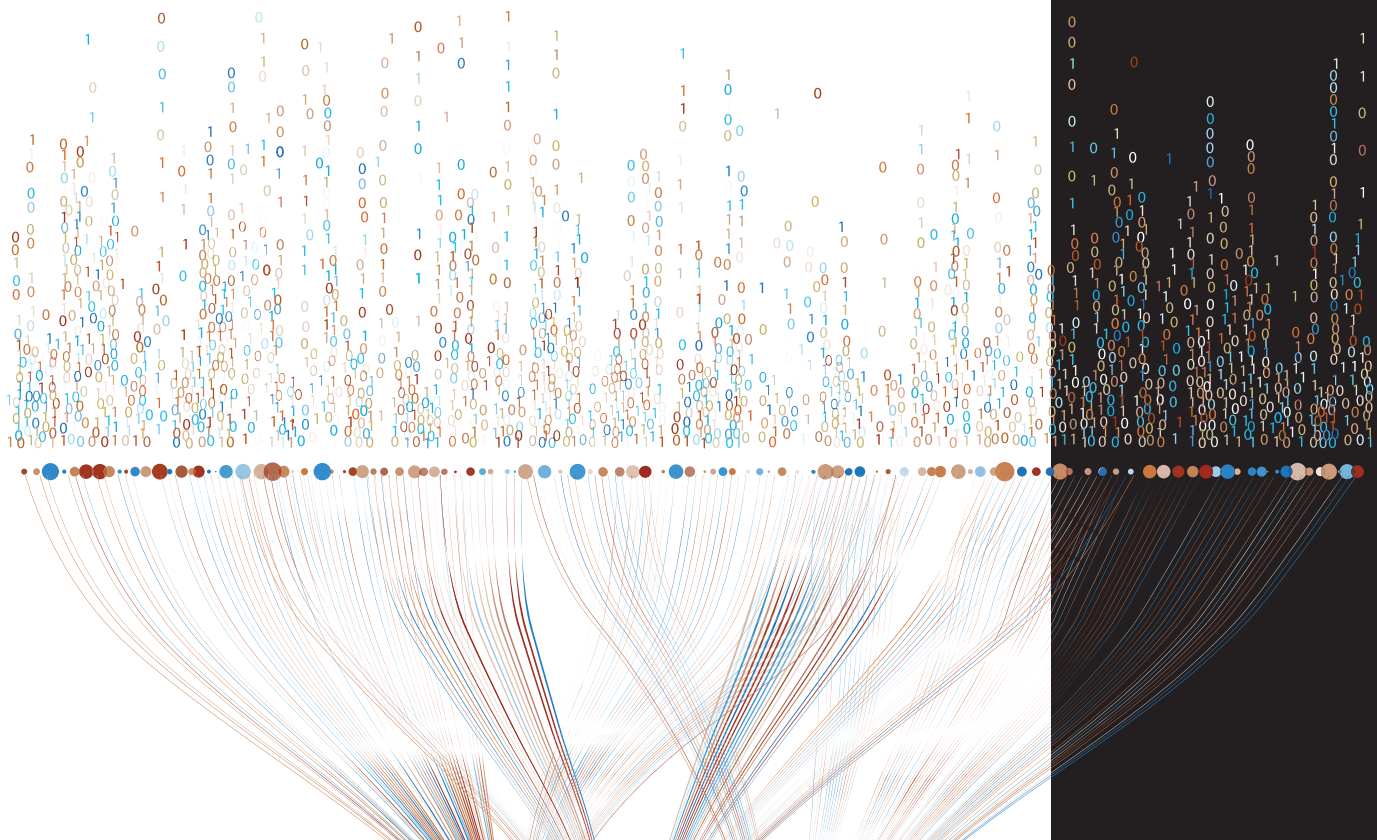
"There are a lot of these companies that are conglomerating and buying up stuff and trying to build this. But that takes so long, it's almost easier for us to look at doing this in-house and getting all that into a data lake, letting it all talk together, identifying the fields and seeing where the commonalities are," said Ryan Johnson, CTO, Rocklin Unified School District.

At El Dorado County, Digiorno is counting on a newly established IT governance model to help simplify integration and security, with data sharing as an overarching goal.

The county decommissioned its mainframe system shortly before Digiorno arrived in late 2019, which has created integration and data challenges the organization didn't face previously.

"When you have a single system, it's easy to control the data, but when they started branching off into commercial off-the-shelf software and some custom development, the integration isn't always on point," she said. For Digiorno, an IT steering committee is the first step toward establishing the organizational coordination needed for data sharing.

"It's a big effort that takes the whole village," she said. "It's not an IT director-driven effort; everybody needs to be a part of it."





# Analytics Challenge 4: Finding Talent and Skills

**Our CIO Think Tank panelists validated** a huge pain point that companies everywhere encounter: Good data scientists are hard to find. So are data architects and engineers. The analytics arms race has created huge demand for these professionals, outstripping the pool of available candidates.

"The talent, global talent, U.S. talent, it is not there," was the blunt assessment of Wolters Kluwer's Gangavarapu.

Different kinds of businesses reported various constraints.

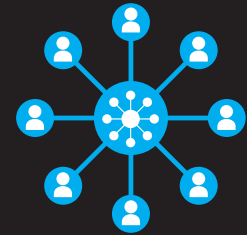
"As a school district, we face that jack-of-all-trades, master-of-none idea, where we have some data-base guys, and we've got a little bit of data analytics, but we don't have people that are dedicated to that," said Johnson.

Across the board, panelists said hiring is extremely difficult.

Other, subtler dimensions to the skills challenge persist. A data scientist skilled in building models doesn't necessarily have the business context to understand the data to be analyzed. Likewise, data experts might not have the storytelling skills to communicate conclusions to a diverse audience of stakeholders.

Also, as analytics programs grow and mature, the skills needed will change. That complicates both hiring and training.

"We thought about actually hiring somebody and bringing them in-house, but there is certain expertise that you want to have at the table while you initiate the project, and then it may turn into a different expertise that you need for an ongoing maintenance and execution," said Meyer.



# 46%

Rise in data-science hiring  
between 2019 and 2020,  
according to LinkedIn

Source: Jobs on the Rise



## RECOMMENDATIONS:

Our panelists suggested a wide variety of tactics while acknowledging that the talent challenge is difficult.

First, go local—with universities, nonprofits, and vocational programs to help build your own team. At Replacement Parts Inc., Spugnardi has had success working with a group in Little Rock called the Arkansas Center for Data Sciences to bring in apprentices and hire entry-level employees.

"They pay for the training. I give them jobs. It works out well," he said, while also noting that his team must supply training in the business context.

Others have found more candidates by broadening their search around the world.

"We've been challenged [to hire] in India where it felt like we are not necessarily the employer of choice. There's a lot of competition there," said Richard Thompson, managing director, BNY Mellon. "We ventured into Poland and found an uptick there in terms of talent and availability."

It's nearly inevitable for CIOs to lean on vendor partners for staffing projects and to train internal employees as the work moves along.

"I think that's where we can probably use the most leaning in from vendors—the augmentation of resources," said Digiorno.

Cool tools help attract good technologists. This is an age-old IT conundrum: Businesses shouldn't adopt technology for its own sake. However, said HPE's Matt Maccaux, "They don't want irrelevant skills on their resume, and they want to be as 'cloud' as possible." He advocates for the use of open systems wherever possible. "That's very attractive for developers, data scientists, and machine learning engineers to want to build on top of my software platform," he said.

Training is certainly part of the solution; many panelists said they are implementing data-focused training programs within IT, in addition to their data literacy efforts across the business.

"People get motivated when they get a new opportunity, so one thing we do is move people around [and] try to position people based on their interests and how they'd like to move," said Gangavarapu.

Lastly, the dearth of talent may influence organizational structure, suggesting a central team early in the journey to take best advantage of scarce experts.

"The human resources constraints with data scientists and that community are driving us to a shared services model. That has kind of forced us into some of the natural centralization and automation," said Sandesh Shetty, CTO for Koch.

# Analytics Challenge 5: Democratizing Analytics

**Scant benefit lies in having lots of information** unless many people are using it. That imperative has given rise to the “citizen data scientist,” Eric Knorr said, attributing the phrase to InfoWorld contributor [Isaac Sacolick](#).

This is not to say that the average line of business employees will become expert in neural networks. Sacolick said there is a class of problems that will always require a high degree of expertise, but a much larger set of problems that data or business analysts can address with visualization and reporting tools. Citizen data scientists will combine business domain knowledge with enough data skills to answer business questions and to create dashboards and other tools to help colleagues do the same themselves.

In an organizational sense, data analytics is not just about scaling up, but scaling out to the people who can use it in everyday operations. That's the way organizations will ultimately achieve maximum insights and business value—with dozens, hundreds, or even thousands of employees making data-informed decisions and products.

CIO Think Tank panelists noted that this will require even more new tools, skills, and processes that are, in some cases, fundamentally different from those employed in earlier stages of the analytics journey.

Notably, of the 31 participants across three panel discussions, no executive claimed that their organization has arrived at this desired state.

Ultimately, democratized analytics will require mirroring the cultural values and qualities you've worked so hard to create (see section 1) in workflows, processes, and systems. That will take time, and good prioritization will be the key to success.

“Where’s the biggest ROA, or return on automation, if you will? That’s where you want to be,” said Crowley.



**\$215.7  
BILLION**

**Expected in global spending on big data and business analytics (BDA) solutions in 2021** Source: IDC



## RECOMMENDATIONS:

**Always keep an eye toward self-service, said Jyoti. "Democratizing these tools and technologies is very, very important—how we empower the business analysts so that they're not shy of visualizing and dicing up this data," he explained.**

**If there is an end state in the analytics journey, this is it: End users across the organization have the skills and tools to make data-informed decisions without waiting for help from a bottlenecked department of experts.**

**Many of the solutions or approaches suggested in preceding challenges can play a role in this democratization.**

**A more data-literate workforce is readier to use tools and data as they become available. A strong, centralized data organization can provide literacy training and help with the tools as the workforce ramps up.**

**"In the early stages, you form a centralized team for all the business units to use. Eventually, as the skillset becomes more prevalent, you have the teams move out into the business units," said HPE's Goh.**

**For self-service, many companies have added tools that allow users to rummage through data without needing to write SQL queries or use R for statistical programming. However, additional and easy-to-use tools will be a must.**

**"There is no way this industry is going to scale without having some kind of no-code or low-code version of AI or ML analytics to hit these tens of thousands of narrow-use cases that people want to solve problems for," said Robert Christiansen, VP Strategy, HPE.**

**Maccaux's advocacy of open systems and standards applies as well, with an eye toward using widely accessible systems and minimizing technical debt while adopting new tools. Examples of such software include the popular self-service business intelligence tools Tableau and PowerBI.**

**"As long as you're using open-source standards and open APIs for your data access layer—your data cataloging layer—that's going to allow whomever you hire to use the tools of their choice and get onto the path of monetizing the data," he said.**

## **IDG**

**Eric Knorr**  
B2B Editor in Chief  
Eric\_Knorr@IDG.com

## **IDC**

**Ritu Jyoti**  
Group VP, Worldwide AI and  
Automation Research Practice

## **HPE**

**Robert Christiansen**  
VP, Strategy

**Dr. Eng Lim Goh**  
SVP and CTO, High-  
performance Computing and AI

**Matt Maccaux**  
Global Field CTO  
Ezmeral Software, HPE

## **MODERATOR**

**John Gallant**  
IDG Enterprise  
Consulting Director  
john\_gallant@idg.com

**Rick Spugnardi**, CIO, Replacement Parts Inc.

**Keith Jefferson**, VP, Citigroup

**Tonya Digiorno**, IT Director, El Dorado County

**Ryan Johnson**, CTO, Rocklin USD

**Timothy (Tim) J. Dion**, Chief Data Officer, Armor Healthcare

**Shobie Ramakrishnan**, Chief Digital and Technology Officer,  
Pharma Commercial, GlaxoSmithKline

**Sunil Badiani**, VP, Cyber Security Officer, BNY Mellon

**Subbu Ramanathan**, Head of Software and Web Solutions Development  
(Divisional CIO), Brady Corp.

**Susan Bagh**, VP Strategy and Transformation, Wells Fargo

**Stephen Held**, VP/CIO, LEO A DALY

**Sandesh Shetty**, CTO, Koch

**Sandeep Sidhu**, CIO, Emily Carr University

**Dave Rogers**, Senior Director, BI, Raising Cane's Chicken Finger

**Bob Lim**, VP IT and CIO, San Jose State University

**Jason See**, CIO, Raymond W Bliss Army Health Center

**Adam Garey**, VP Operations Services and Application Development, Country Financial

**Paul Groll**, Director of Emerging Technology Research, State of Michigan

**Dierdre (Dee) Bonora**, CIO, Orrstown Bank

**Raja Gangavarapu**, VP, Wolters Kluwer

**Pavi Agrawal**, Chief Applications Architect, Mars

**Richard Thompson**, Managing Director, BNY Mellon

**Nazmul Islam**, Head of IT, UAB Department of Medicine

**Rebecca Meyer**, IS Director, Commerce Apps and Ecommerce, Kelly Moore Paint

**Alexander (Alex) Mendoza**, Director, Data and Analytics, Chobani LLC

**Gurpreet Atwal**, Senior VP, Mastercard

**Michael Crowley**, VP, Technology and Innovation, Martin-Brower Co.

**Steve Baumgartner**, CIO, Inpro Corporation

**James Song**, VP, IT, Blue Cross Blue Shield of North Carolina

**Mike Wasik**, CIO, Flagstone Foods

**Manish Desai**, Cyber Security Senior Information Risk Officer, BNY Mellon

**Gayathri Mariappan**, IT Director, RLI