THE RISE OF ANALYTICS 3.0

How to Compete in the Data Economy

By Thomas H. Davenport, IIA Research Director

INTERNATIONAL
INSTITUTE FOR
ANALYTICS℠

About IIA

IIA is an independent research firm for organizations committed to accelerating their business through the power of analytics. We believe that in the new data economy only those who compete on analytics win. We know analytics inside and out—it's what we do. IIA works across a breadth of industries to uncover actionable insights gleaned directly from our network of analytics practitioners, industry experts, and faculty. In the era of analytics, we are teachers, guides and advisors. The result? Our clients learn how best to leverage the power of analytics for greater success in the new data economy.

Let IIA be your guide. Contact us to accelerate your progress:

Click: iianalytics.com
Call: 503-467-0210
Email: research@iianalytics.com

About Thomas Davenport

IIA Research Director Tom Davenport is the President's Distinguished Professor of IT and Management at Babson College, and a research fellow at the MIT Center for Digital Business. Tom’s "Competing on Analytics" idea was recently named by *Harvard Business Review* as one of the twelve most important management ideas of the past decade and the related article was named one of the ten ‘must read’ articles in HBR’s 75 year history. His most recent book, co-authored with Jinho Kim, is *Keeping Up with the Quants: Your Guide to Understanding and Using Analytics.*

Analytics 3.0 is an environment that combines the best of 1.0 and 2.0—a blend of big data and traditional analytics that yields insights with speed and impact.

CONTENTS

Introduction:
The Rise of Analytics 3.0

Big-thinking gurus often argue that we have moved from the agricultural economy to the industrial economy to the data economy.

It is certainly true that more and more of our economy is coordinated through data and information systems. However, outside of the information and software industry itself, it's only over the last decade or so that data-based products and services really started to take off.

The impact of the data economy goes beyond previous roles for data and analytics; instead of slowly improving back-office decisions, the data economy promises new business models and revenue sources, entirely new operational and decision processes, and a dramatically accelerated timescale for business.

Around the turn of the 21st century, as the Internet became an important consumer and business resource, online firms including Google, Yahoo, eBay, and Amazon began to create new products and services for customers based on data and analytics.

A little later in the decade, online network-oriented firms including Facebook and LinkedIn offered services and features based on network data and analytics.

These firms created the technologies and management approaches we now call "big data," but they also demonstrated how to participate in the data economy. In those companies, data and analytics are not just an adjunct to the business, but the business itself.

THE PIONEERS

Examples of products and services based on data and analytics:

Google
search algorithms

amazon.com
product recommendations

ebay
fraud detection

facebook
"Custom Audiences" for targeted ads

LinkedIn
"People You May Know"

As large firms in other industries also adopt big data and integrate it with their previous approaches to data management and analytics, they too can begin to develop products and services based on data and analytics, and enter the data economy.

GE is one of the early adopters of this approach among industrial firms. It is placing sensors in "things that spin" such as jet engines, gas turbines, and locomotives, and redesigning service approaches based on the resulting data and analysis.

Since half of GE's revenues in these businesses come from services, the data economy becomes critical to GE's success.

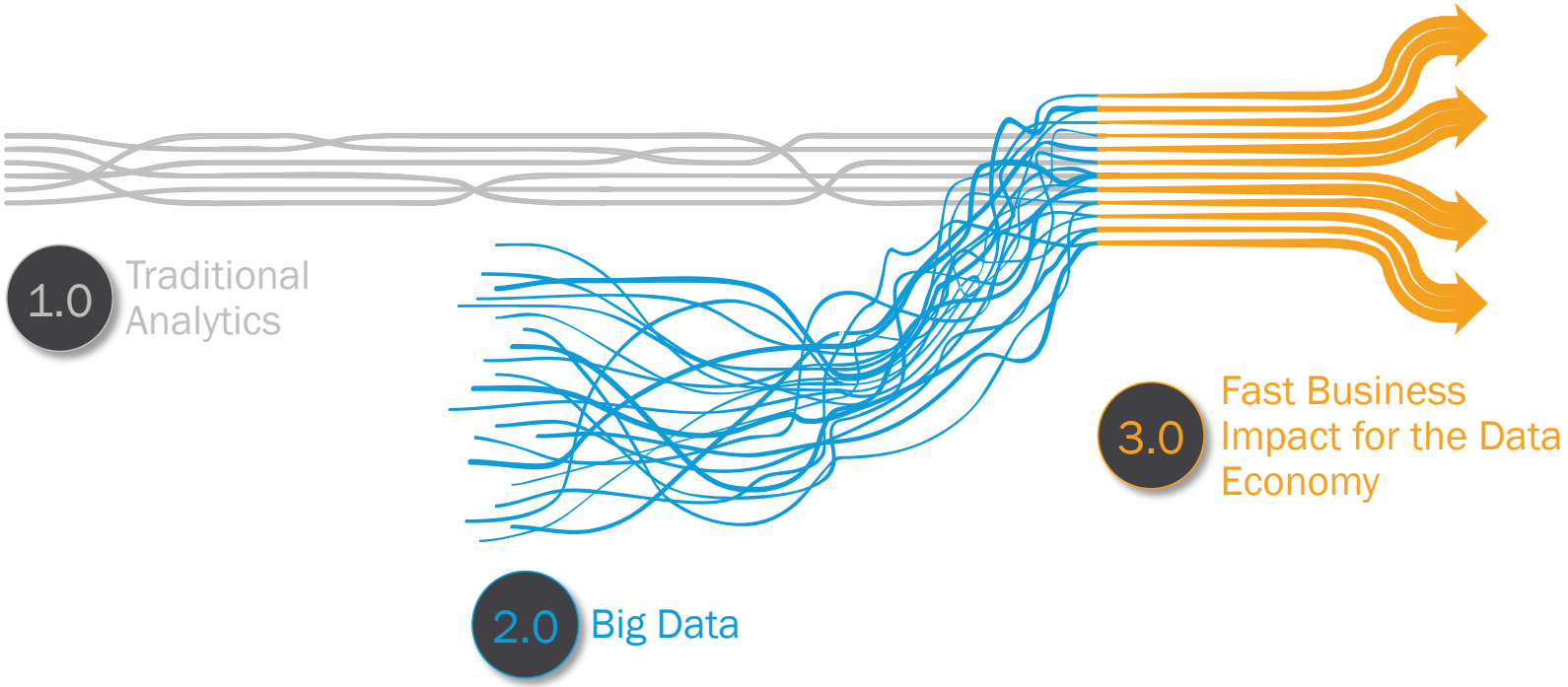
Analytics 3.0 in Action

"things that spin"

How GE adopted a big data approach:

- $2B initiative in software and analytics
- Primary focus on data-based products and services from "things that spin"
- Will reshape service agreements for locomotives, jet engines, and turbines
- Gas blade monitoring in turbines produces 588 gigabytes/day—7 times Twitter daily volume

ANALYTICS 3.0 | FAST BUSINESS IMPACT FOR THE DATA ECONOMY



- "Analytics 3.0" is an appropriate name for this next evolution of the analytics environment, since it follows upon two earlier generations of analytics use within organizations.
- Analytics 3.0 represents a new set of opportunities presented by the data economy, and a new set of management approaches at the intersection of traditional analytics and big data.

Analytics 1.0—Traditional Analytics

Analytics are not a new idea. To be sure, there has been a recent explosion of interest in the topic, but for the first half-century of activity, the way analytics were pursued in most organizations didn't change much.

This Analytics 1.0 period predominated for half a century from the mid-1950s (when UPS initiated the first corporate analytics group in the U.S.) to the mid-2000s, when online firms began innovating with data.

Of course, firms in traditional offline industries continued with Analytics 1.0 approaches for a longer period, and many organizations still employ them today.

Analytics 1.0 Characteristics

1.0 Traditional Analytics

- Data sources relatively small and structured, from internal systems;
- Majority of analytical activity was descriptive analytics, or reporting;
- Creating analytical models was a time-consuming "batch" process;
- Quantitative analysts were in "back rooms" segregated from business people and decisions;
- Few organizations "competed on analytics"—analytics were marginal to strategy;
- Decisions were made based on experience and intuition.

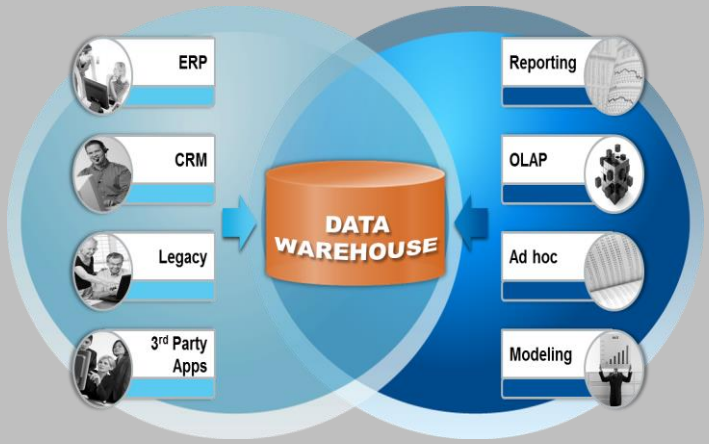
From a technology perspective, this was the era of the enterprise data warehouse and the data mart. Data was small enough in volume to be segregated in separate locations for analysis.

This approach was successful, and many enterprise data warehouses became uncomfortably large because of the number of data sets contained in them. However, preparing an individual data set for inclusion in a warehouse was difficult, requiring a complex ETL (extract, transform, and load) process.

For data analysis, most organizations used proprietary BI and analytics "packages" that had a number of functions from which to select.

Analytics 1.0 Data Environment

This was the era of the enterprise data warehouse and the data mart.



More than 90% of the analysis activity involved descriptive analytics, or some form of reporting.

The Analytics 1.0 ethos was internal, painstaking, backward-looking, and slow. Data was drawn primarily from internal transaction systems, and addressed well-understood domains like customer and product information.

Reporting processes only focused on the past, without explanation or prediction. Statistical analyses often required weeks or months. Relationships between analysts and decision-makers were often distant, meaning that analytical results often didn't meet executives' requirements, and decisions were made on experience and intuition.

Analysts spent much of their time preparing data for analysis, and relatively little time on the quantitative analysis itself.

Analytics 1.0 Ethos

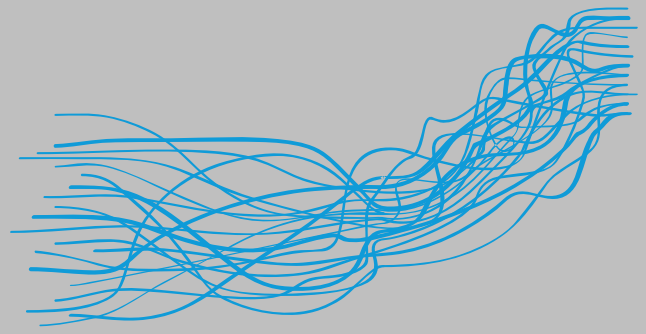
- Stay in the back room—as far away from decision-makers as possible—and don't cause trouble
- Take your time—nobody's that interested in your results anyway
- Talk about "BI for the masses," but make it all too difficult for anyone but experts to use
- Look backwards—that's where the threats to your business are
- If possible, spend much more time getting data ready for analysis than actually analyzing it
- Keep inside the sheltering confines of the IT organization

Analytics 2.0—Big Data

Starting in the mid-2000s, the world began to take notice of big data (though the term only came into vogue around 2010), and this period marked the beginning of the *Analytics 2.0* era.

The period began with the exploitation of online data in Internet-based and social network firms, both of which involved massive amounts of fast-moving data. Big data and analytics in those firms not only informed internal decisions in those organizations, but also formed the basis for customer-facing products, services, and features.

Analytics 2.0 Characteristics



2.0 Big Data

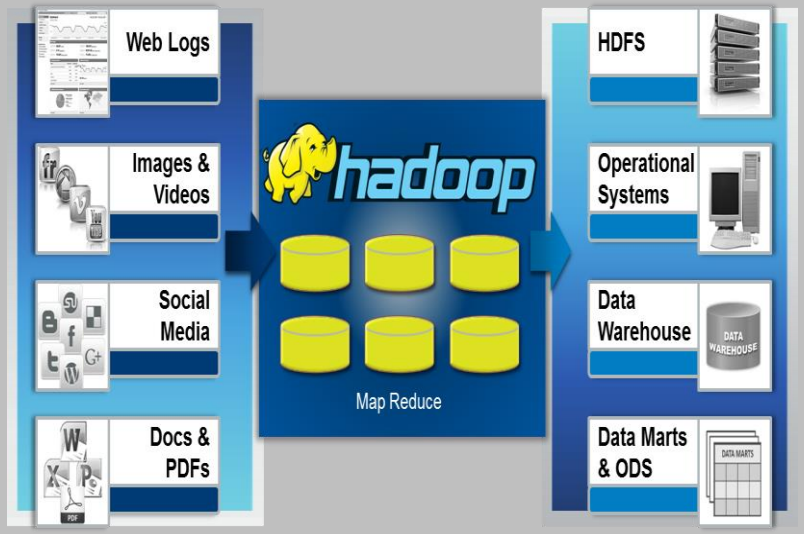
- Complex, large, unstructured data sources
- New analytical and computational capabilities
- "Data Scientists" emerge
- Online firms create data-based products and services

These pioneering Internet and social network companies were built around big data from the beginning. They didn't have to reconcile or integrate big data with more traditional sources of data and the analytics performed upon them, because for the most part they didn't have those traditional forms.

They didn't have to merge big data technologies with their traditional IT infrastructures because those infrastructures didn't exist.

Big data could stand alone, big data analytics could be the only focus of analytics, and big data technology architectures could be the only architecture.

Analytics 2.0 Data Environment



Big data analytics as a standalone entity in Analytics 2.0 were quite different from the 1.0 era in many ways.

As the big data term suggests, the data itself was either very large, relatively unstructured, fast-moving—or possessing all of these attributes.

Data was often externally-sourced, coming from the Internet, the human genome, sensors of various types, or voice and video.

A new set of technologies began to be employed at this time.

Key Developments in 2.0

- Fast flow of data necessitated rapid storage and processing
- Parallel servers running Hadoop for fast batch data processing
- Unstructured data required "NoSQL" databases
- Data stored and analyzed in public or private cloud computing environments
- "In-memory" analytics and "in-database" analytics employed
- Machine learning methods meant the overall speed of analysis was much faster (from days to minutes)
- Visual analytics often crowded out predictive and prescriptive techniques

The ethos of Analytics 2.0 was quite different from 1.0. The new generation of quantitative analysts was called “data scientists,” with both computational and analytical skills.

Many data scientists were not content with working in the back room; they wanted to work on new product offerings and to help shape the business.

There was a high degree of impatience; one big data startup CEO said, “We tried agile [development methods], but it was too slow.” The big data industry was viewed as a “land grab” and companies sought to acquire customers and capabilities very quickly.

Analytics 2.0 Ethos

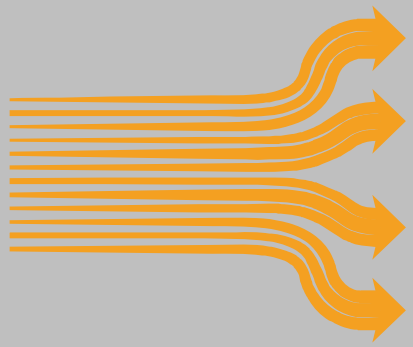
- Be “on the bridge” if not in charge of it
- “Agile is too slow”
- “Being a consultant is the dead zone”
- Develop products, not PowerPoints or reports
- Information (and hardware and software) wants to be free
- All problems can be solved in a hackathon
- Share your big data tools with the community
- “Nobody’s ever done this before!”

Analytics 3.0—Fast Impact for the Data Economy

Big data is still a popular concept, and one might think that we're still in the 2.0 period. However, there is considerable evidence that large organizations are entering the Analytics 3.0 era.

It's an environment that combines the best of 1.0 and 2.0—a blend of big data and traditional analytics that yields insights and offerings with speed and impact.

Analytics 3.0 Characteristics

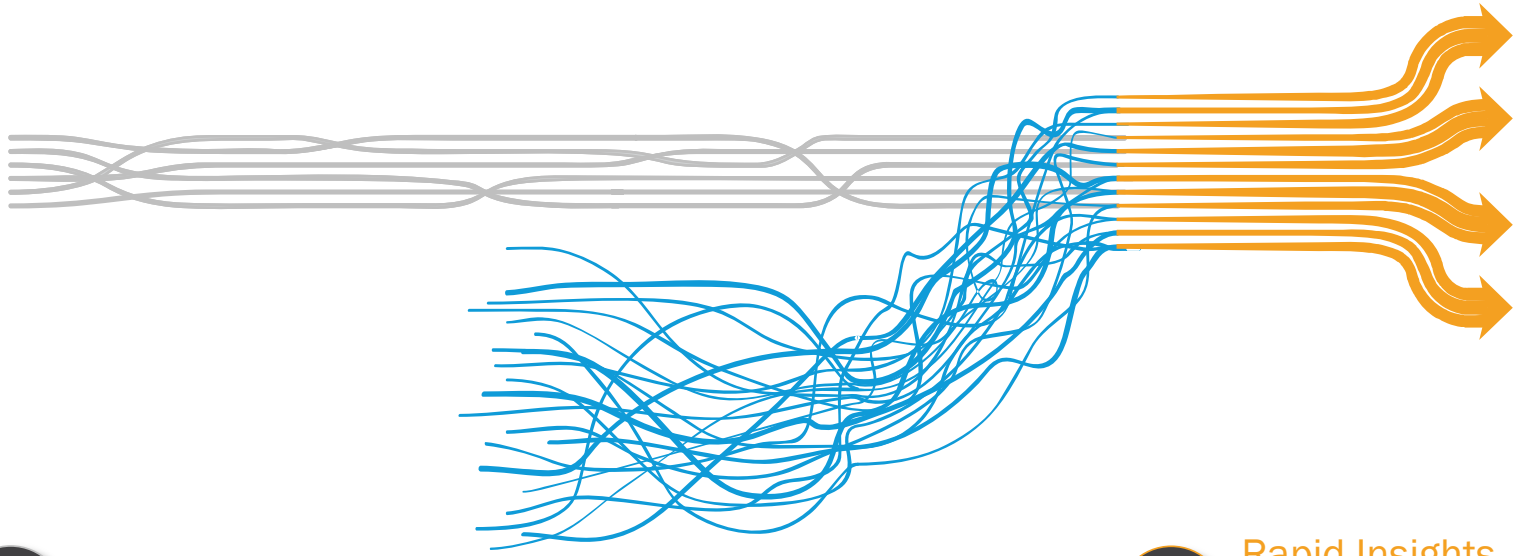


3.0 Rapid Insights Providing Business Impact

- Analytics integral to running the business; strategic asset
- Rapid and agile insight delivery
- Analytical tools available at point of decision
- Cultural evolution embeds analytics into decision and operational processes
- All businesses can create data-based products and services

ANALYTICS 3.0 | THE ERA OF IMPACT

INTERNATIONAL INSTITUTE FOR ANALYTICS



1.0 Traditional Analytics

- Primarily descriptive analytics and reporting
- Internally sourced, relatively small, structured data
- "Back room" teams of analysts
- Internal decision support

2.0 Big Data

- Complex, large, unstructured data sources
- New analytical and computational capabilities
- "Data Scientists" emerge
- Online firms create data-based products and services

3.0 Rapid Insights Providing Business Impact

- Analytics integral to running the business; strategic asset
- Rapid and agile insight delivery
- Analytical tools available at point of decision
- Cultural evolution embeds analytics into decision and operational processes
- All businesses can create data-based products and services

Although it's early days for this new model, the traits of Analytics 3.0 are already becoming apparent.

Banks, industrial manufacturers, health care providers, retailers—any company in any industry that is willing to exploit the possibilities—can all develop data-based offerings for customers, as well as support internal decisions with big data.

The most important trait of the Analytics 3.0 era is that not only online firms, but virtually any type of firm in any industry, can participate in the data economy.

There is considerable evidence that when big data is employed by large organizations, it is not viewed as a separate resource from traditional data and analytics, but merged together with them.

In addition to this integration of the 1.0 and 2.0 environments, other attributes of Analytics 3.0 organizations are described on the following pages.

Analytics 3.0 in Action

"From the beginning of our Science function at AIG, our focus was on both traditional analytics and big data. We make use of structured and unstructured data, open source and traditional analytics tools. We're working on traditional insurance analytics issues like pricing optimization, and some exotic big data problems in collaboration with MIT. It was and will continue to be an integrated approach."

—Murli Buluswar, Chief Science Officer
AIG

Multiple Data Types, Often Combined

Organizations are combining large and small volumes of data, internal and external sources, and structured and unstructured formats to yield new insights in predictive and prescriptive models.

Often the increased number of data sources is viewed as incremental, rather than a revolutionary advance in capability.

Analytics 3.0 in Action

At Schneider National, a large trucking firm, the company is increasingly adding data from new sensors—monitoring fuel levels, container location and capacity, driver behavior, and other key indicators—to its logistical optimization algorithms.

The goal is to improve—slowly and steadily—the efficiency of the company's route network, to lower the cost of fuel, and to decrease the risk of accidents.

A New Set of Data Management Options

In the 1.0 era, firms employed data warehouses with copies of operational data as the basis for analysis. In the 2.0 era, the focus was on Hadoop clusters and NoSQL databases.

Now, however, there are a variety of options from which to choose in addition to these earlier tools: Database and big data appliances, SQL-to-Hadoop environments (sometimes called "Hadoop 2.0"), vertical and graph databases, etc.

Enterprise data warehouses are still very much in evidence as well. The complexity and number of choices that IT architects have to make about data management have expanded considerably, and almost every organization will end up with a hybrid data environment.

The old formats haven't gone away, but new processes need to be developed by which data and the focal point for analysis will move across staging, evaluation, exploration, and production applications.

Technologies and Methods are Much Faster

Big data technologies from the Analytics 2.0 period are considerably faster than previous generations of technology for data management and analysis.

To complement the faster technologies, new "agile" analytical methods and machine learning techniques are being employed that produce insights at a much faster rate.

Like agile system development, these methods involve frequent delivery of partial outputs to project stakeholders; as with the best data scientists' work, there is an ongoing sense of urgency.

The challenge in the 3.0 era is to adapt operational and decision processes to take advantage of what the new technologies and methods can bring forth.

Integrated and Embedded Analytics

Consistent with the increased speed of analytics and data processing, models in Analytics 3.0 are often being embedded into operational and decision processes, dramatically increasing their speed and impact.

Some firms are embedding analytics into fully automated systems based on scoring algorithms or analytics-based rules. Others are building analytics into consumer-oriented products and features.

In any case, embedding the analytics into systems and processes not only means greater speed, but also makes it more difficult for decision-makers to avoid using analytics—usually a good thing.

Analytics 3.0 in Action

P&G

Procter & Gamble's leadership is passionate about analytics and has moved business intelligence from the periphery of operations to the center of how business gets done.

P&G embeds analytics in day-to-day management decision-making, with its "Business Sphere" management decision rooms and over 50,000 desktops equipped with "Decision Cockpits."

Hybrid Technology Environments

It's clear that the Analytics 3.0 environment involves new technology architectures, but it's a hybrid of well-understood and emerging tools.

The existing technology environment for large organizations is not being disbanded; some firms still make effective use of relational databases on IBM mainframes.

However, there is a greater use of big data technologies like Hadoop on commodity server clusters; cloud technologies (private and public), and open-source software.

Analytics 3.0 Data Environment

The most notable changes in the 3.0 environment are new options for data storage and analysis , and attempts to eliminate the ETL (extract, transform, and load) step before data can be assessed and analyzed.

This objective is being addressed through real-time messaging and computation tools such as Apache Kafka and Storm.

Data Science/Analytics/IT Teams

Data scientists often are able to run the whole show—or at least have a lot of independence—in online firms and big data startups. In more conventional large firms, however, they have to collaborate with a variety of other players.

In many cases the "data scientists" in large firms may be conventional quantitative analysts who are forced to spend a bit more time than they like on data management activities (which is hardly a new phenomenon).

And the data hackers who excel at extracting and structuring data are working with conventional quantitative analysts who excel at modeling it.

This collaboration is necessary to ensure that big data is matched by big analytics in the 3.0 era.

Both groups have to work with IT, which supplies the big data and analytical infrastructure, provides the “sandboxes” in which they can explore data, and who turns exploratory analyses into production capabilities.

Teams of data hackers and quant analysts are doing whatever is necessary to get the analytical job done, and there is often a lot of overlap across roles.

Chief Analytics Officers

When analytics and data become this important, they need senior management oversight.

And it wouldn't make sense for companies to have multiple leaders for different types of data, so they are beginning to create "Chief Analytics Officer" roles or equivalent titles to oversee the building of analytical capabilities. We will undoubtedly see more such roles in the near future.

Organizations with C-level analytics roles:

- AIG
- University of Pittsburgh Medical Center
- The 2012 Barack Obama reelection campaign
- Wells Fargo
- Bank of America
- USAA
- RedBox
- FICO
- Teradata

The Rise of Prescriptive Analytics

There have always been three types of analytics:

1. **descriptive**, which report on the past;
2. **predictive**, which use models based on past data to predict the future;
3. **prescriptive**, which use models to specify optimal behaviors and actions.

Analytics 3.0 includes all types, but there is an increased emphasis on prescriptive analytics.

These models involve large-scale testing and optimization. They are a means of embedding analytics into key processes and employee behaviors. They provide a high level of operational benefits for organizations, but require high-quality planning and execution.

Analytics 3.0 in Action

UPS is using data from digital maps and telematics devices in its trucks to change the way it routes its deliveries.

The new system (called ORION, for On-Road Integrated Optimization and Navigation) provides routing information to UPS's 55,000 drivers.

Summary

Even though it hasn't been long since the advent of big data, these attributes add up to a new era.

It is clear from our research that large organizations across industries are joining the data economy. They are not keeping traditional analytics and big data separate, but are combining them to form a new synthesis.

Some aspects of Analytics 3.0 will no doubt continue to emerge, but organizations need to begin transitioning now to the new model.

It means changes in skills, leadership, organizational structures, technologies, and architectures. Together these new approaches constitute perhaps the most sweeping change in what we do to get value from data since the 1980s.

Analytics 3.0

Competing In The Data Economy

- Every company—not just online firms—can create data and analytics-based products and services that change the game
- Not just supplying data, but customer insights and guides to decision-making
- Use "data exhaust" to help customers use your products and services more effectively
- Start with data opportunities or start with business problems? Answer is yes!
- Need "data products" team good at data science, customer knowledge, new product/service development
- Opportunities and data come at high speed, so quants must respond quickly

It's important to remember that the primary value from big data comes not from the data in its raw form, but from the processing and analysis of it and the insights, decisions, products, and services that emerge from analysis.

The sweeping changes in big data technologies and management approaches need to be accompanied by similarly dramatic shifts in how data supports decisions and product/service innovation processes.

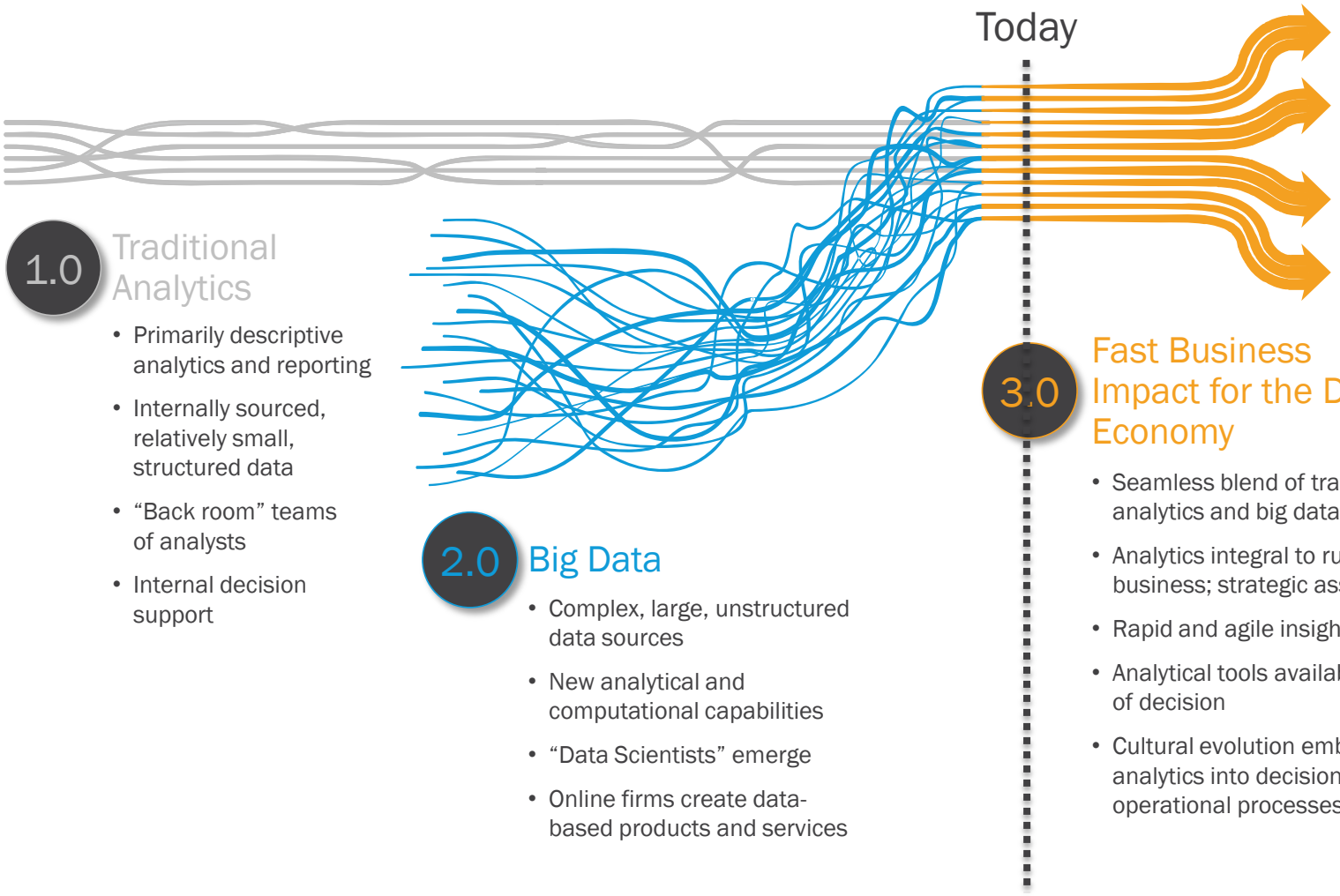
These shifts have only begun to emerge, and will be the most difficult work of the Analytics 3.0 era. However, there is little doubt that analytics can transform organizations, and the firms that lead the 3.0 charge will seize the most value.

The primary value from big data comes not from the data in its raw form, but from the processing and analysis of it and the insights, decisions, products, and services that emerge from analysis.

SUMMARY OF THE IDEA

Era	1.0: Traditional Analytics	2.0: Big Data	3.0: Data Economy
Timeframe	Mid-1950s to 2000	Early 2000s to Today	Today and in the Future
Culture, Ethos	Very few firms "compete on analytics"…"we know what we know."	Agile, experimental, hacking…new focus on data-based products and services for customers	Agile methods that speed "time to decision"…all decisions driven (or influenced) by data…the data economy
Type of analytics	• 5% predictive, prescriptive • 95% reporting, descriptive	• 5% predictive, prescriptive • 95% reporting, descriptive (visual)	• 90%+ predictive, prescriptive • Reporting automated commodity
Cycle time	Months ("batch" activity)	An insight a week	Millions of insights per second
Data	Internal, structured…very few external sources available or perceived as valuable **"data"**	Very large, unstructured, multi-source…much of what's interesting is external…explosion of sensor data **"Big Data"**	Seamless combination of internal and external…analytics embedded in operational and decision processes… tools available at the point of decision **"Data Economy"**
Technology	Rudimentary BI, reporting tools…dashboards…data stored in enterprise data warehouses or marts	New technologies: Hadoop, commodity servers, in-memory, machine learning, open source…"unlimited" compute power	New data architectures…beyond the warehouse New application architectures…specific apps, mobile
Organization & Talent	Analytical people segregated from business and IT…"Back Room" statisticians, quants without formal roles	Data Scientists are "on the bridge"…talent shortage noted…educational programs on the rise	Centralized teams, specialized functions among team members, dedicated funding… Chief Analytics Officers … recognized training, education programs

ANALYTICS 3.0 | FAST BUSINESS IMPACT FOR THE DATA ECONOMY



1.0 Traditional Analytics

- Primarily descriptive analytics and reporting
- Internally sourced, relatively small, structured data
- "Back room" teams of analysts
- Internal decision support

2.0 Big Data

- Complex, large, unstructured data sources
- New analytical and computational capabilities
- "Data Scientists" emerge
- Online firms create data-based products and services

3.0 Fast Business Impact for the Data Economy

- Seamless blend of traditional analytics and big data
- Analytics integral to running the business; strategic asset
- Rapid and agile insight delivery
- Analytical tools available at point of decision
- Cultural evolution embeds analytics into decision and operational processes

Recipe for a 3.0 World

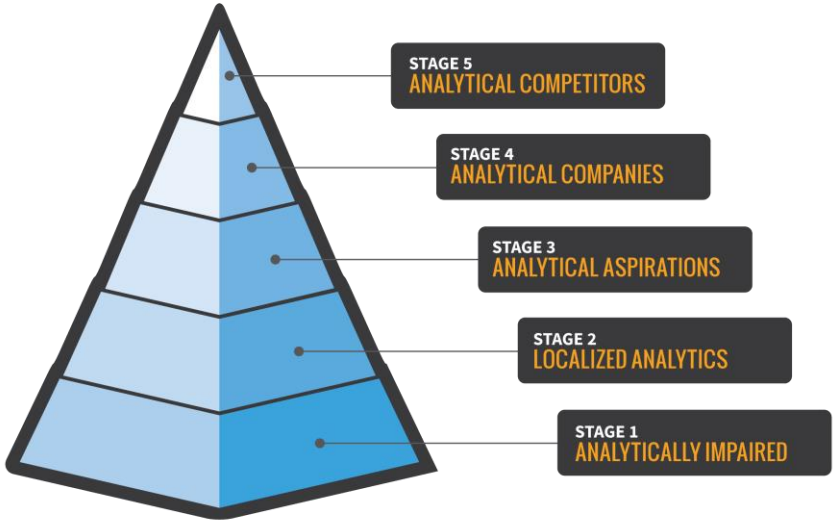
- Start with an existing capability for data management and analytics
- Add some unstructured, large-volume data
- Throw some product/service innovation into the mix
- Add a dash of Hadoop
- Cook up some data in a high-heat convection oven
- Embed this dish into a well-balanced meal of processes and systems
- Promote the chef to Chief Analytics Officer

Recommendations for Success in the New Data Economy

- Use five factors (DELTA), five levels (1-5) to establish and measure analytical capability
- Become a student of the analytics environment
- Educate your leaders on the potential of analytics based on what you're seeing
- Use IIA as your barometer

Δ DELTA = CHANGE

D	DATA	BREADTH, INTEGRATION, QUALITY
E	ENTERPRISE	APPROACH TO MANAGING ANALYTICS
L	LEADERSHIP	PASSION AND COMMITMENT
T	TARGETS	FIRST DEEP THEN BROAD
A	ANALYSTS	PROFESSIONALS AND AMATEURS



ARE YOU POISED TO COMPETE IN THE DATA ECONOMY?

Let IIA be your guide.

Contact us to accelerate your progress.

Click: iianalytics.com

Call: 503-467-0210

Email: research@iianalytics.com

INTERNATIONAL INSTITUTE FOR ANALYTICS™