TODAY'S AGENDA

- David Stodder, TDWI

- Ben Smith, HPE Vertica

- Tony Ordner & Steven John, AmeriPride

- Roundtable discussion

- Closing Q&A & Wrap Up

Developing a Strategy for Advancing Faster with Big Data Analytics

DAVID STODDER

Senior Research Director,
Business Intelligence
TDWI

MY AGENDA

- Introduction and trends
- Role of data warehouse modernization
- TDWI's Big Data Maturity Model
- Maturity strategies and recommendations

Getting from Here to There: Strategies for Increasing Maturity with Big Data Analytics

Why Big Data Analytics? Business Disruption

- Need to respond to increased, sometimes unexpected competition

- Track changes in customer behavior

- Rapid (or slowing) growth detected in core markets; need to know why

- Company want to deploy a new business model or enter a new market
  - Repeating the "same old" formula is not an option

- Leadership change: Top executives want to be more data- and analytics-driven

- Regulatory changes

- Because it's there (the data, that is)
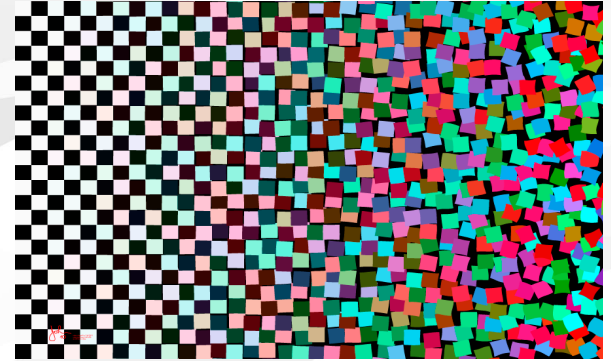
Credit: www.john-art.com

Big Data *Plus* Analytics: Pushing the Envelope

- **Big data**: High-volume, structured, semi-, and unstructured data
  - Not just volume: variety, velocity
  - Internal and external sources
  - New data types; e.g. geospatial
  - Streaming, machine, IoT data

- **Beyond the warehouse**: Adding data lakes, cloud, MPP, Hadoop/MapReduce
  - Revolutionary Hadoop ecosystem now 11 years old!

- **Data science**: Moving analytics beyond BI, OLAP, and traditional statistics
  - Experimenting with data; looking for correlations
  - AI and machine learning

Analytics Vision: Smarter Response to Disruption

- **Power of analytics**: Upper management is paying attention
  - "Competing on analytics" is top of mind; disruptive analytics for new business models
  - Good for the data-savvy: Creating a path for data-savvy professionals to influence leadership

- **To be data-informed**: Basing decisions on analytics
  - What are the risks and rewards of a decision? Operational impact? Prediction and scenarios
  - "Citizen data science": democratization of analytics and less dependence on specialists
  - Visual analytics: New language of collaboration

- **Faster innovation**: Using analytics to shorten decision cycles
  - Goal: Higher efficiency and speed in making and acting on decisions; push for real-time analytics
  - Detecting patterns, anomalies, and changes sooner

Big Data Analytics: Numerous Use Cases

- Analyzing customer behavior
- Marketing personalization
- Customer segmentation
- Social media analysis
- Fraud detection
- Risk analysis
- Loan approvals
- Facility/inventory monitoring
- Preventive maintenance
- IT operations analytics

Example: Manufacturing, Moving from Process- to Analytics-Driven

- Sensors generate big data: IoT enabling new perspectives on processes and potential for optimization
  - Boeing: 8-10k sensors capturing 8k data points per second, from 5k planes
  - Key role of predictive analytics to measure what *will* happen in terms of maintenance
  - Real-time analytics for fuel management and choosing routes

- Using analytics to correct process flaws sooner
  - Product quality and defects tracking
  - Supply planning
  - Output forecasting
  - Energy efficiency
  - Mass customization

TDWI Research: Many Organizations are Collecting Big Data Volumes

How much data does your organization collect and use for analytics?
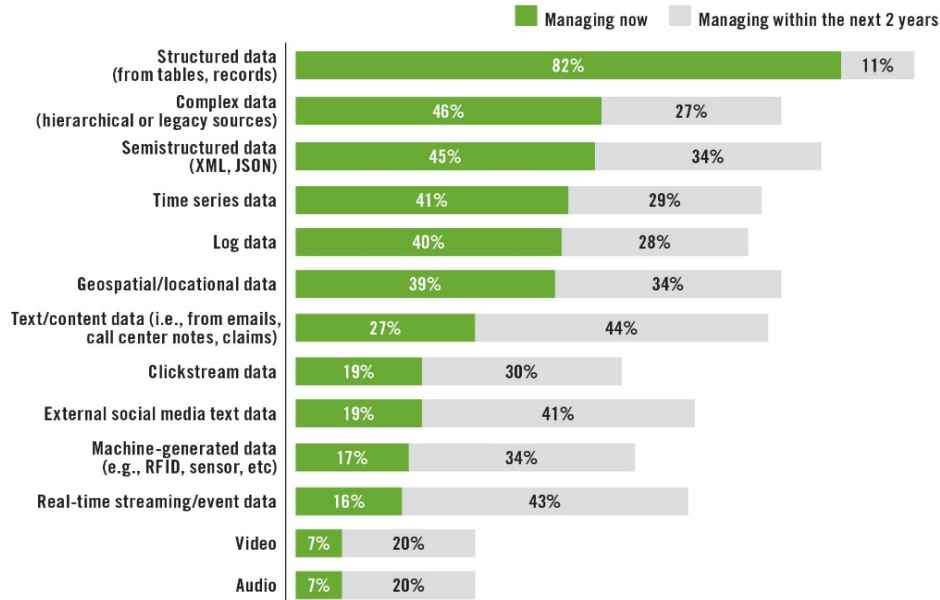

- Don't know 7%
- Megabytes 2%
- 7% Over 1 petabyte
- 10% 100–1,000 terabytes
- Gigabytes 24%
- 20% 10–100 terabytes
- Less than 10 terabytes 30%

37%
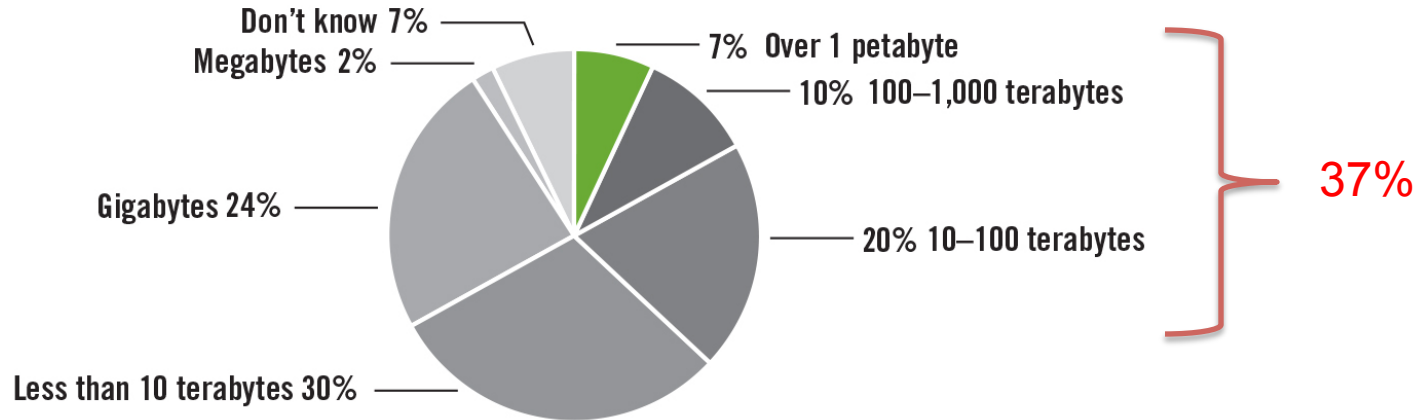
Figure 2. Based on 370 respondents.

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

However, the Data is Mostly Structured

Which of the following types of data are you managing as big data now? Within two years from now?

Legend: ■ Managing now ■ Managing within the next 2 years

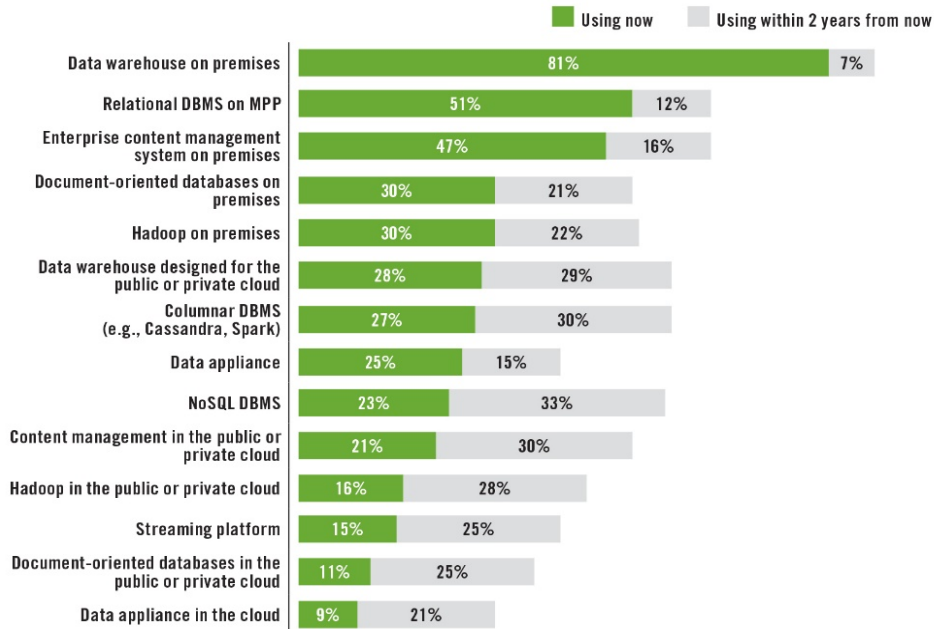| Data type	Managing now	Managing within the next 2 years
Structured data (from tables, records)	82%	11%
Complex data (hierarchical or legacy sources)	46%	27%
Semistructured data (XML, JSON)	45%	34%
Time series data	41%	29%
Log data	40%	28%
Geospatial/locational data	39%	34%
Text/content data (i.e., from emails, call center notes, claims)	27%	44%
Clickstream data	19%	30%
External social media text data	19%	41%
Machine-generated data (e.g., RFID, sensor, etc)	17%	34%
Real-time streaming/event data	16%	43%
Video	7%	20%
Audio	7%	20%

Figure 3. Ordered by percentage managing each data type now. Remainder don't know or have no plans. Based on 370 respondents.

Source: TDWI Best Practices Report on Big Data and Data Science, 2016

tdwi **TRANSFORMING DATA WITH INTELLIGENCE**™

…And On a Range of Platforms

What kind of data management platforms are you using for big data now? Within two years from now?

■ Using now   ■ Using within 2 years from now

Platform	Using now	Using within 2 years from now
Data warehouse on premises	81%	7%
Relational DBMS on MPP	51%	12%
Enterprise content management system on premises	47%	16%
Document-oriented databases on premises	30%	21%
Hadoop on premises	30%	22%
Data warehouse designed for the public or private cloud	28%	29%
Columnar DBMS (e.g., Cassandra, Spark)	27%	30%
Data appliance	25%	15%
NoSQL DBMS	23%	33%
Content management in the public or private cloud	21%	30%
Hadoop in the public or private cloud	16%	28%
Streaming platform	15%	25%
Document-oriented databases in the public or private cloud	11%	25%
Data appliance in the cloud	9%	21%

Figure 5. *Ordered by percentage using each data platform now. Remainder don't know or have no plans. Based on 338 respondents.*

Source: TDWI Best Practices Report on Big Data and Data Science, 2016

DW Modernization: From DW Only to Multiple Data Platforms

Which of the following best describes the architecture of your extended DW environment today? What about in three years?
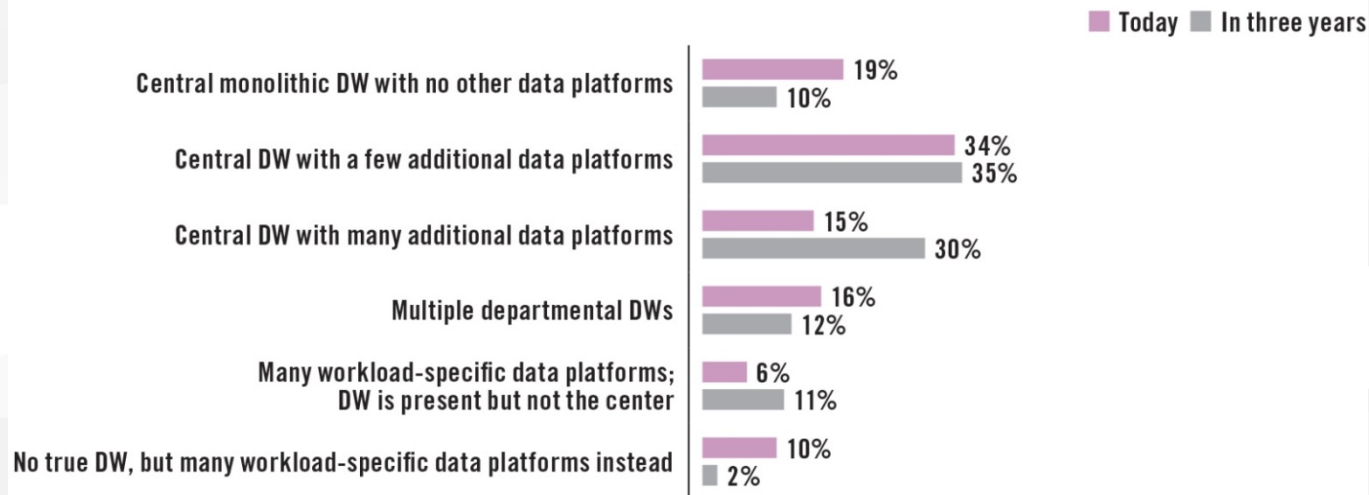
Legend: ■ Today ■ In three years

Architecture	Today	In three years
Central monolithic DW with no other data platforms	19%	10%
Central DW with a few additional data platforms	34%	35%
Central DW with many additional data platforms	15%	30%
Multiple departmental DWs	16%	12%
Many workload-specific data platforms; DW is present but not the center	6%	11%
No true DW, but many workload-specific data platforms instead	10%	2%

Figure 15. *Based on 473 respondents.*

tdwi **TRANSFORMING DATA WITH INTELLIGENCE**™

What Is Data Warehouse Modernization?

- To support new technology and business requirements, organizations are
  - Augmenting
  - Automating
  - Optimizing

  …their data warehouse

- Data warehouse environments are becoming "hybrid"
  - Adding purpose-built analytics platforms (e.g., columnar, graph, NoSQL, virtual)

- Technology advancements
  - In-memory and in-database analytics
  - Massively parallel processing
  - Fresher data, streaming, and real-time operations

- Hadoop and data lakes
  - Growing interest and maturity; use for ETL, staging area

Biggest Barriers to DW Modernization

In your organization, what are the top barriers to data warehouse modernization? (Respondents could provide multiple answers)

- Poor stewardship or governance of data: 40%

- Inadequate staffing for data warehousing and related disciplines: 39%

- Cost of implementing new technologies: 34%

- Inadequate skills for new technologies and practices: 33%

- Lack of business case or sponsorship: 30%

- Lack of experience with new big data types and their analytics: 28%

Source: TDWI Best Practices Report on Data Warehouse Modernization, 2016

tdwi **TRANSFORMING DATA WITH INTELLIGENCE**™

DW Modernization and Big Data Analytics

- **Analytics**: One of the top drivers for DW modernization, according to TDWI Research
  - Need to support projects for data mining, predictive analytics, and more

- Analytics agility in a single environment:
  - Need to support different styles of analytics: batch, interactive, and real time
  - Hybrid environments

- **New workloads**: Matching workloads with right platforms
  - Business-driven: What is the purpose and priority?
  - Run on the DW? The analytic platform? The Hadoop data lake (which might also need modernizing)?
  - "Structure" isn't a bad thing: Data lakes can become "swamps" where it's hard to do anything

- **The Cloud**: Growing as a data source and processing option

Data Modernization for Big Data Analytics: Recommendations

- ✓ Realign the environment with business goals
  - "If your project aligns with your business goals, your entire business will modernize, not just your warehouse. After all, that's the real point." – Philip Russom, TDWI

- ✓ Assume need for multiple manifestations of modernization
  - System, non-DW, continuous, disruptive; need multiple strategies

- ✓ Consider Hadoop ecosystem technologies for various roles
  - Data staging, ELT pushdown, archiving of detailed source data (retained for analytics)

- ✓ DW modernization is not "one and done"
  - Develop plans for recurring cycles

Strategies for Increasing
Maturity with Big Data Analytics

TDWI Big Data Maturity Model

Model and Assessment Goal: Help organizations understand how their big data and analytics deployments compared with those of peers, and how they could develop a road map for advancement

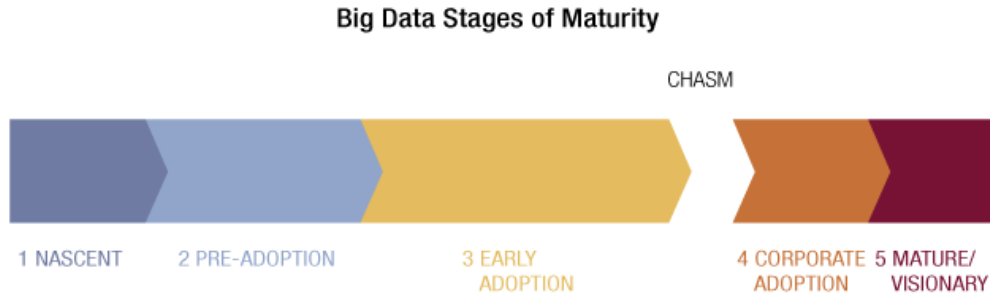**TDWI BIG DATA MATURITY MODEL AND ASSESSMENT TOOL**

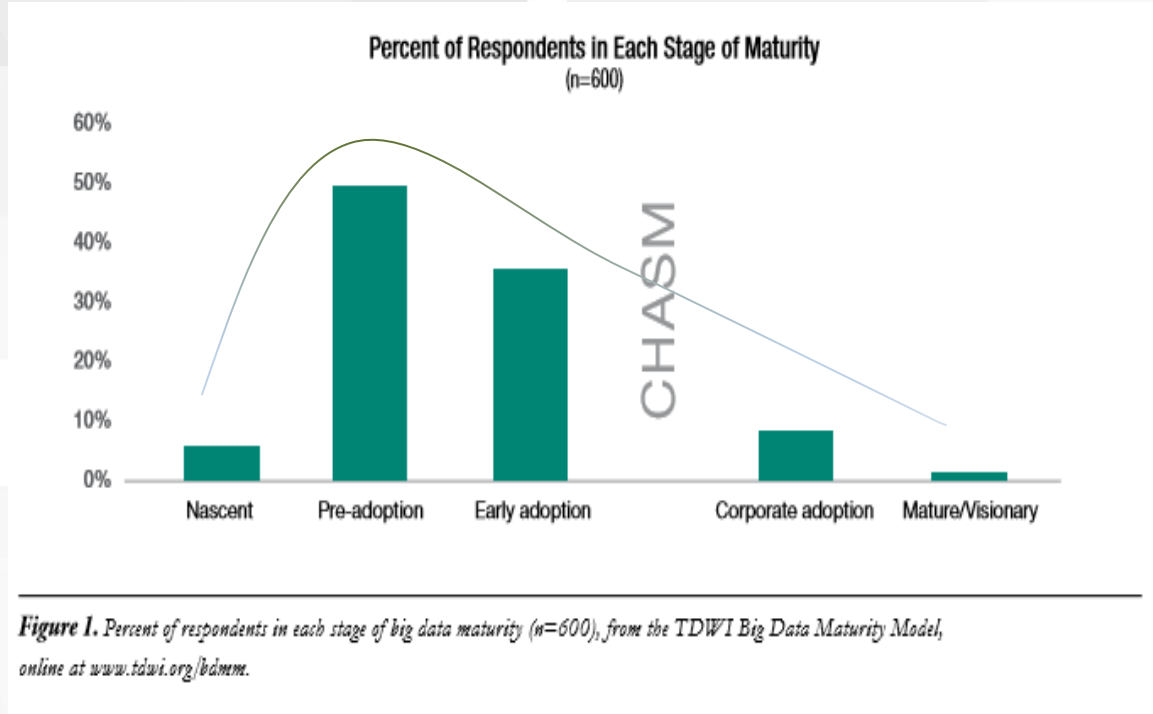**How mature is your enterprise's big data program?**

Take The Assessment

www.tdwi.org (research and resources)

Maturity Assessment

- Model consists of five stages

- Respondents are evaluated against five dimensions: organization, infrastructure, data management, analytics, and governance
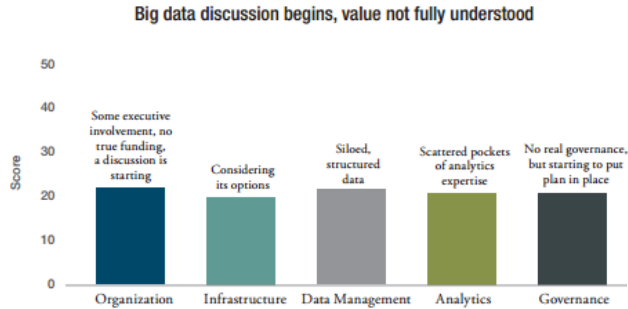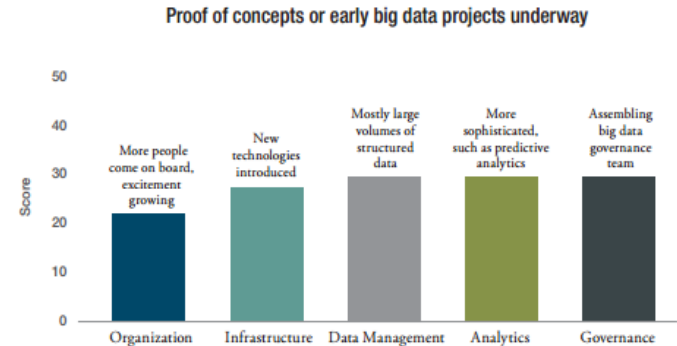
- Online tool provides benchmark

Big Data Stages of Maturity

CHASM

1 NASCENT   2 PRE-ADOPTION   3 EARLY ADOPTION   4 CORPORATE ADOPTION 5 MATURE/ VISIONARY

TDWI Assessment: Results

Percent of Respondents in Each Stage of Maturity
(n=600)

Figure 1. *Percent of respondents in each stage of big data maturity (n=600), from the TDWI Big Data Maturity Model, online at www.tdwi.org/bdmm.*

Characteristics of these companies

Stage 2: Pre-adoption

Big data discussion begins, value not fully understood

Stage 3: Early Adoption

Proof of concepts or early big data projects underway

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

Early Adoption

Cross the Chasm

Corporate Adoption

GETTING FROM HERE TO THERE

tdwi TRANSFORMING DATA WITH **INTELLIGENCE**™

Stage 3: Early Adoption

Proof of concepts or early big data projects underway

Bar chart with Score (y-axis 0–50):
- Organization: ~22 — More people come on board, excitement growing
- Infrastructure: ~27 — New technologies introduced
- Data Management: ~29 — Mostly large volumes of structured data
- Analytics: ~29 — More sophisticated, such as predictive analytics
- Governance: ~29 — Assembling big data governance team

Stage 4: Corporate Adoption

Big data analytics transforms the business

Bar chart with Score (y-axis 0–50):
- Organization: ~39 — Big data seen as strategic, processes changing
- Infrastructure: ~36 — Start of unified architecture
- Data Management: ~41 — Utilizing disparate data
- Analytics: ~41 — Advanced analytics, analytics operationalized
- Governance: ~39 — Governance council in place, policies documented

Seven Strategies for Gaining Maturity

- Gaining Business Buy-in Strategy

- Organizational Strategy

- Analytics Strategy

- Data Management Strategy

- Infrastructure Strategy

- Governance Strategy

- A "Visionary" Strategy

STRATEGY #1

GAINING BUSINESS BUY-IN

Business Buy-In

- Garner executive support

- Use POCs to show the difference

- Evangelize, socialize, be persistent

- Collaborate and communicate

Resistance   Engagement
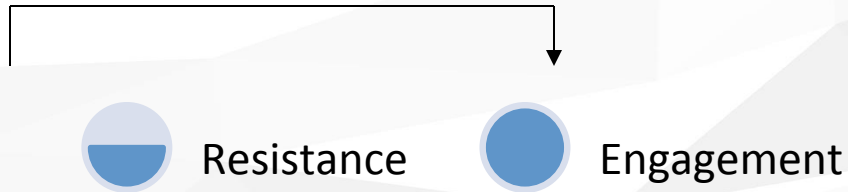
Value of Executive Support

- Can set tone and vision; provide funding
- Aligns with leadership of CIO, VP Analytics, Chief Data, Analytics Officers
- Need to set expectations

STRATEGY #2

ORGANIZATIONAL STRATEGY

TRANSFORMING DATA WITH **INTELLIGENCE**™

Organizational Strategies

- Build talent/skill sets
- Create a Center of Excellence
- Build deployment teams

Skills Needed in the Team
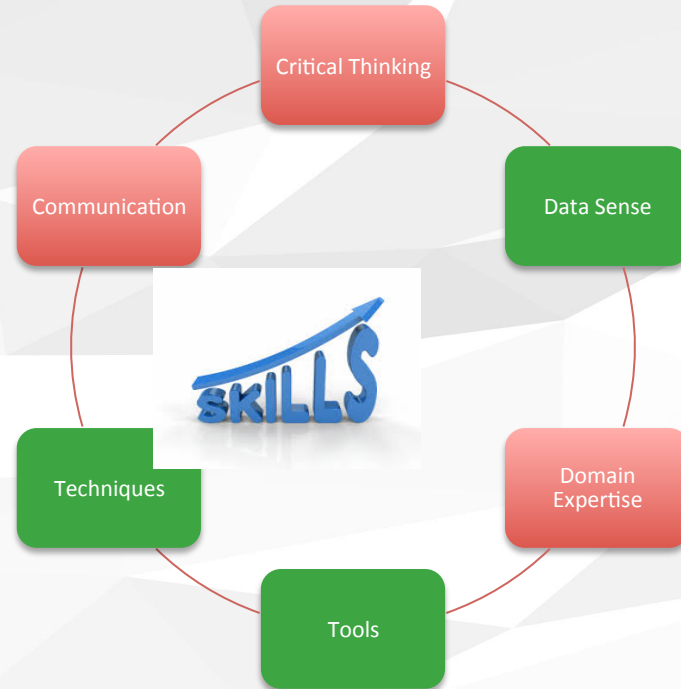
Center of Excellence Strategies

Centralized
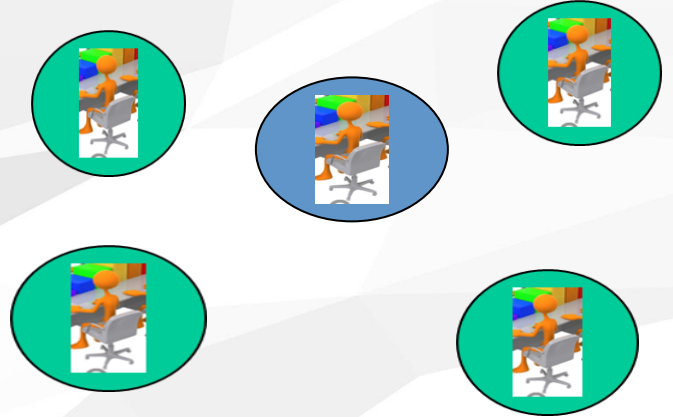
Distributed

STRATEGY #3

ANALYTICS STRATEGY

Analytics Strategies

- Expand the data mix

- Utilize multiple analytics techniques
  - Move beyond descriptive and predictive analytics to include prescriptive
  - Consider the cloud

- Operationalize and embed analytics
  - Goal of actionable analytics

STRATEGY #4

DATA MANAGEMENT STRATEGY

Mature by Expanding Management to New Data Types and Platforms

- Prepare for semi- and unstructured data
  - Customer behavior, social media, JSON, text sources
  - Streaming, IoT

- Evaluate technologies
  - Data lake management
  - Analytics platforms
  - Data integration and preparation tools
  - Cloud options

- Get ready for analytics of detailed data
  - OLAP-style aggregates, samples, etc., are still important, but big data analytics projects often need to run against huge volumes of detailed data
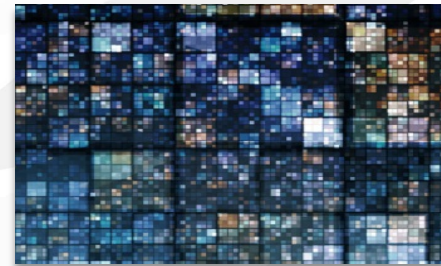
STRATEGY #5

INFRASTRUCTURE STRATEGY

Getting Ready for Increase Everywhere

- **Thinking growth**: Big data analytics success will breed more and larger projects

- Infrastructure concerns:
  - Performance and availability
  - Speed of access and query execution
  - Concurrency
  - Scalability for more users and analytics workloads – for both experimentation and operational decisions

- **Looking to the clouds**: As part of updating their infrastructure strategy, organizations should consider cloud computing and outsourcing

Key Objectives for Gaining Maturity

- **Unified architecture**: Aim to avoid chaos of multiple, diverse data silos
  - Big data analytics requires access to all relevant data
  - Consider integration from the beginning
  - Understand analytics lifecycles; determine appropriate platforms for the workloads
  - Where to put maintenance resources
  - Where to build or contract for skills

- **Think speed**: As projects move beyond "science experiments" and begin to support strategic and operational decisions, speed becomes crucial
  - Assess bottlenecks
  - Consider technologies that support faster analytics: e.g., in-memory computing, columnar databases
  - "Hybrid" architectures that include cloud to "get big fast"

STRATEGY #6

GOVERNANCE STRATEGY

Big Data Analytics Governance Maturity

- Evaluate governance policies and rules for data lakes and cloud sandboxes
  - Common: Role-based access that assigns rights and accountability

- Solidify governance and controls over analytics model deployment
  - Models should be cataloged, managed, and monitored

- Tight, but not too tight: Make sure rules and policies are appropriate and essential
  - Otherwise, users will resort to "shadow" IT systems
  - Bring users and IT together for joint leadership over governance
  - Governance as an enabler

STRATEGY #7

AIM FOR "VISIONARY" LEVEL OF MATURITY

"Visionary": The Top Level of Maturity

- **In the footsteps of the pioneers**: Firms in e-commerce, search, and social networking set the pace in big data analytics

- **Platforms fit for purpose**: Need for "hybrid" architecture that fits together new and established technologies and platforms

- **Visionaries**: Gather those who eat, sleep, and breathe analytics
  - Experimental analytics remain essential, but must now also serve daily decisions
  - Operationalizing analytics heightens need for solid governance

- **Technology's contribution**: Ability to execute on ideas faster and in a shareable and repeatable fashion

Checklist for Reaching Visionary Level

- ✓ Drive continuous innovation through application of big data analytics

- ✓ Increase business agility with big data analytics

- ✓ Develop a data management strategy that delivers detailed, diverse data

- ✓ Modernize the warehouse

- ✓ Unify data architecture and integrate analytics to support smart expansion

- ✓ Improve effectiveness of data governance as big data analytics matures

See "The Road to Becoming a Visionary Big Data Analytics Organization, TDWI Checklist Report 2016 and part of this series.
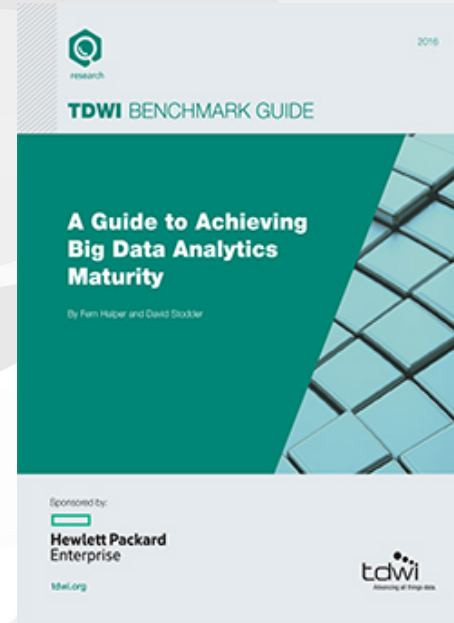
Thank You!

David Stodder
Director of Research for Business Intelligence
TDWI (www.tdwi.org)
dstodder@tdwi.org
(415) 859-9933

Read our TDWI Report

- This presentation includes material from a new report

- Download the free report

- http://bit.ly/2kbjpzY

Learn More in Anaheim!

TDWI Conference

"Modernize. Analyze. Visualize."

Keynotes, Educational Classes, Networking, and More

Anaheim, CA | August 6-11, 2017

http://www.tdwi.org/anaheim

*

TDWI Leadership Summit

"Achieve Faster Business Value with Analytics and BI"

Anaheim, IL | August 7-8, 2017

http://www.tdwi.org/anaheimsummit

tdwi **TRANSFORMING DATA WITH INTELLIGENCE**™