



IBM Predictive Governance

Steven Adler
IBM Information Governance Solutions

Big Data
Predictive Governance
The Impact on Power & Privilege

©ditrak

Agenda

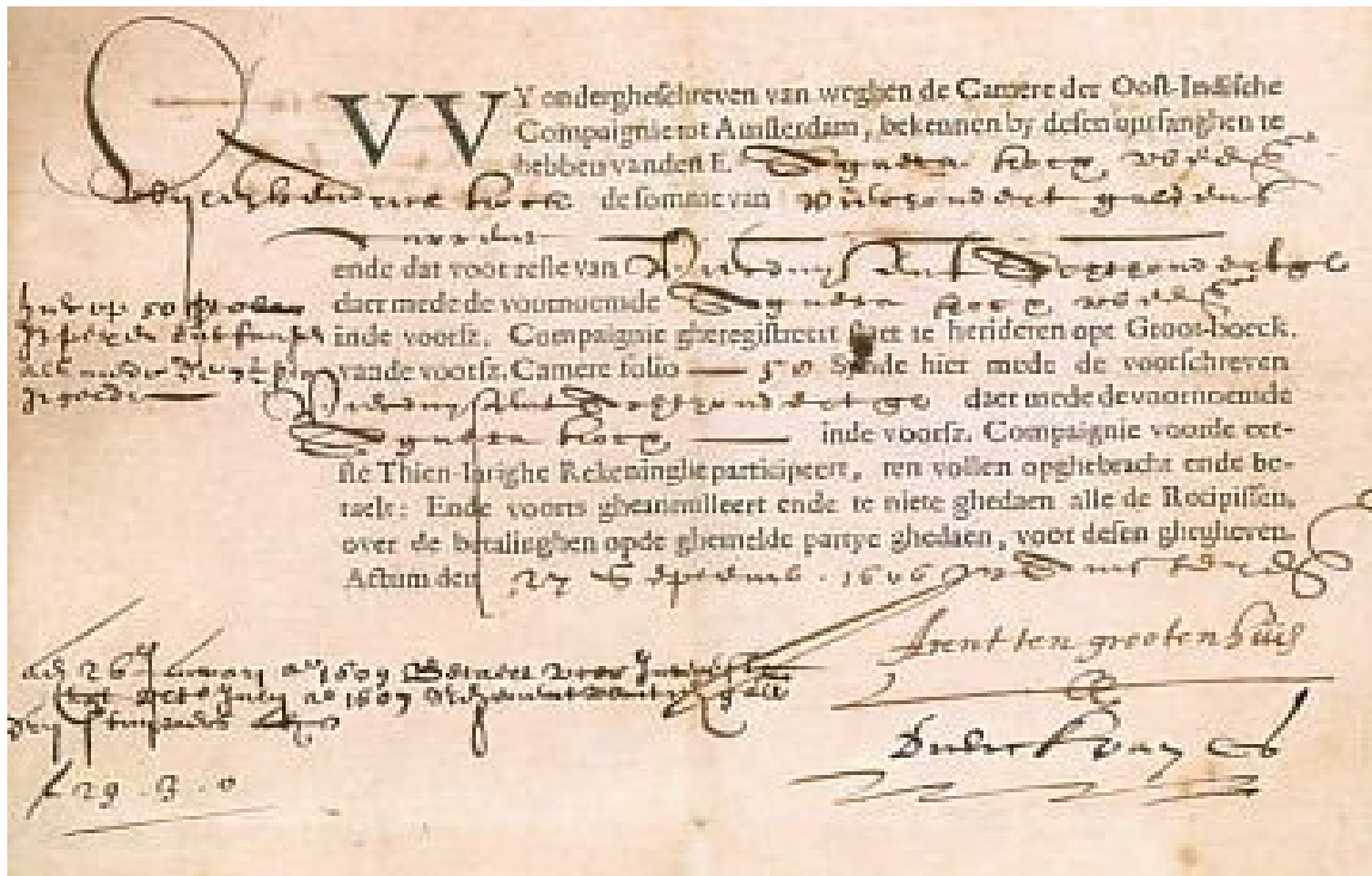
- **Big Data Price Discovery**
- **The Impact on Data Management**
- **Many Challenges**
- **Predictive Governance Simulations**

The First Stock Markets in 1600's...



Were designed for Price Discovery and Risk Management

They were membership clubs...



Whose primary commodity was information

Today markets are electronic...



And membership is still expensive (\$4m) and exclusive

Computers do most of the trading and price discovery

Trade Date: 01/21/2011

NYSE GROUP TRADING INFORMATION	
NYSE Listed	
Matched Share Volume	1,733,486,427
Trades	5,045,343
Value of Trading (\$ millions)	52,489.1
NYSE Arca & AMEX Listed	
Matched Share Volume	256,282,376
Trades	558,857
Value of Trading (\$ millions)	12,424.2
NASDAQ Listed	
Matched Share Volume	196,366,530
Trades	1,041,171
Value of Trading (\$ millions)	7,608.8

NYSE INDEX CLOSES			
Composite	▲	8,105.75	29.03
Energy	▲	12,957.68	93.73
Financial	▲	5,146.94	30.39
Health Care	▲	6,567.66	13.90

NYSE GROUP TRADING INFORMATION YEAR-TO-DATE SUMMARY	
NYSE Listed	
Matched Share Volume (billions of shares)	22.9
Trades (millions of trades)	76.1
Value of Trading (\$ billions)	681.8
NYSE Arca & AMEX Listed	
Matched Share Volume (billions of shares)	4.0
Trades (millions of trades)	10.5
Value of Trading (\$ billions)	188.9
NASDAQ Listed	
Matched Share Volume (billions of shares)	3.5
Trades (millions of trades)	18.3
Value of Trading (\$ billions)	110.6

Number of Trading Days	14
-------------------------------	----

The New York Times

Business Day

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION | A

Search

Global | DealBook | Markets | Economy | Energy | Media | Personal Tech | Small Business | Your Money

Computers That Trade on the News



Minh Uong/The New York Times

By GRAHAM BOWLEY
Published: December 22, 2010

At very high speed using thousands of CPU's

You as an individual trader are seriously...



outclassed by the technology of the market players



Big Data = Volume, Variety and Velocity

Extracting insight from an immense volume, variety and velocity of data, in context, beyond what was previously possible.



Variety: Manage the complexity of multiple relational and unstructured data types and schemas

Velocity: Streaming data and large volume data movement

Volume: Scale from terabytes to zettabytes

Example: A Massive Earthquake in Australia...



Perth

Gold Mine



Will have a ripple effect across South East Asia

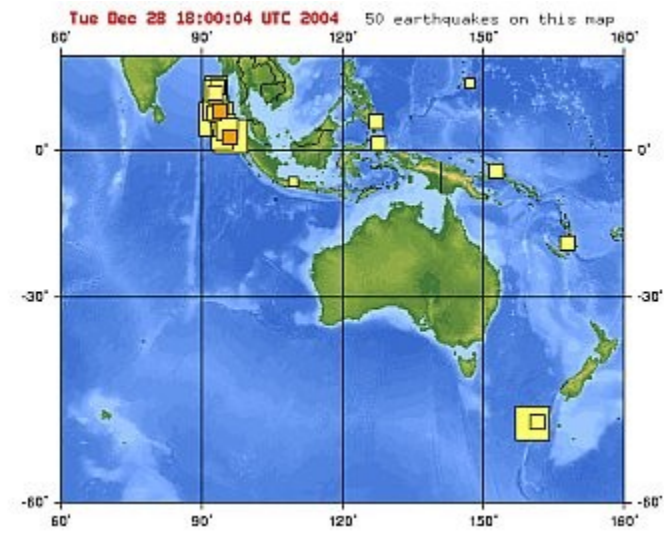
In the old days, markets priced the effects...

Tri Hendri @trihendri 22h
EarthQuake Info: M 5.2, Swan Island, **Australia**: Thursday, May 03, 2012 18:08:06 UTC Thursday, May 03, 2012 12:08:00
on.doi.gov/KgMhQ9
 View location

GOLD MARKET NEWS.EU @GoldMarketNews 30 Apr
#gold #price Early sell shakes **gold** market - Sky News **Australia**
 Early sell shakes **gold** market Sky News **Australia** Early sell s **#Asia**
#news
 Expand

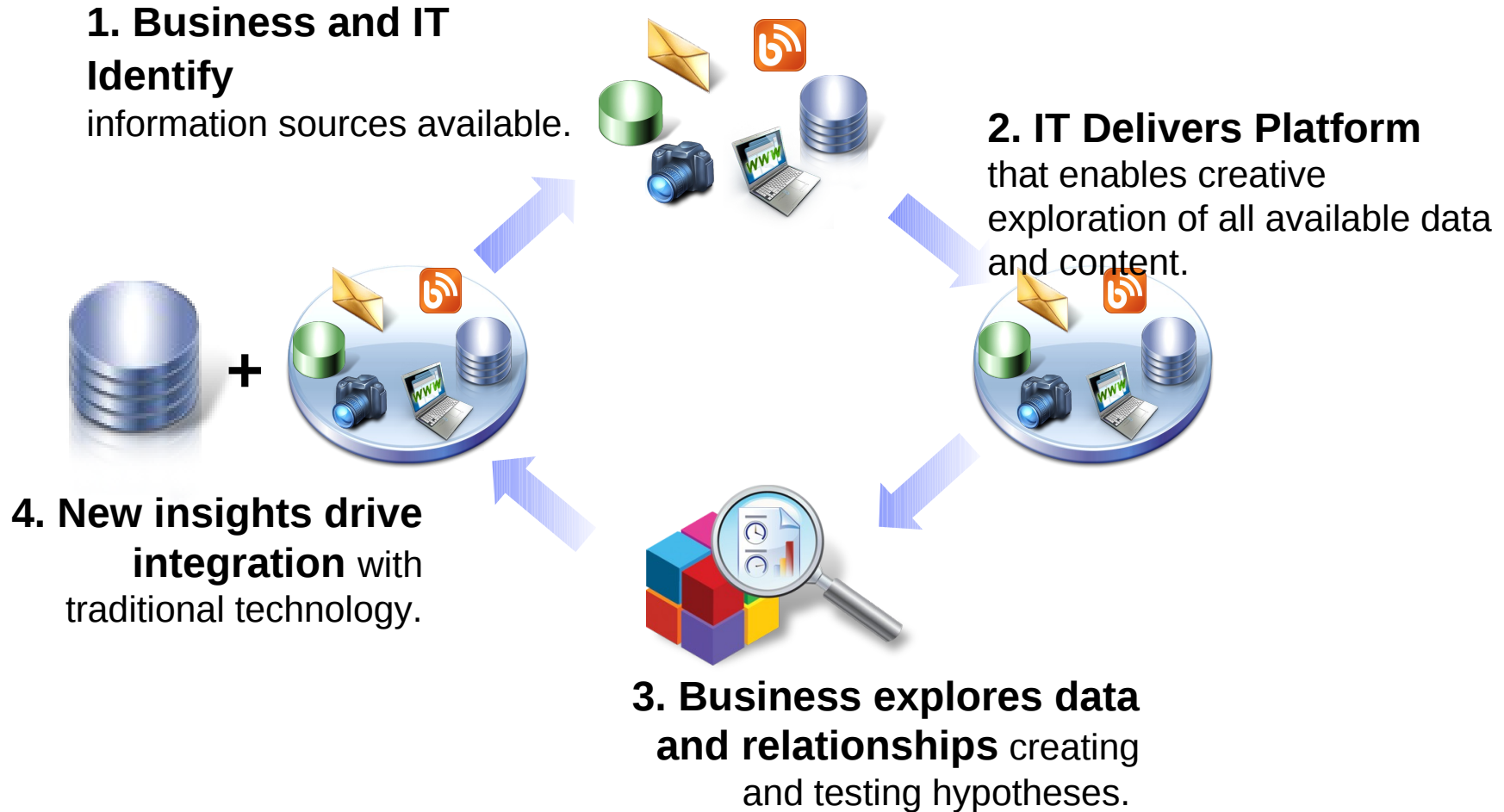
Hard Assets Investor @HAINvestor 28 Apr
 Commodities Week In Review: NatGas Surges 12%, Soybeans **Spike** To 4-Yr High, **Gold** Nears Key Resistance \$SLV \$USO \$CORN hardassetsinvestor.com/weekly-commodi...
 Expand

Shailesh Kumar @arohan 2 May
 Banks Halt Lending to Dry-Bulk **Shipping** Companies, **Genco** Says - Bloomberg ow.ly/aEaAa
 Expand



But Today, you use Big Data to monitor the impact

Big Data is a Paradigm Shift in Data Management



...Data is the Forethought, Applications the Afterthought

Data Governance lifted the profile of DM in 2008...



And Big Data Dramatically Accelerates that Trend

This is just one example – but the list is infinite

Financial Services

- Fraud detection
- 360° View of the Customer



Utilities

- Weather analysis
- Smart grid management



Transportation

- Logistics optimization
- Traffic congestion



I



- System Log Analysis
- Cybersecurity

Health & Life Sciences

- Epidemic early warning
- ICU monitoring



Retail

- 360° View of the Customer
- Real-time promotions

Telecommunications

- Geomapping / marketing
- Network monitoring



Law Enforcement

- Multimodal surveillance
- Cyber security detection

Some More Common Examples

- 1. Use Big Data for Massive Crowdsourcing – adding external expertise to drive product development**
- 2. Perform Natural Language Search to yield insights on the data you have**
- 3. Combine your data with your customers via federation and uncover new insights**
- 4. Mine your employees documents, videos, and emails to automatically update skills in your directory**
- 5. Analyze internal business processes to discover new ways to optimize and cut costs**

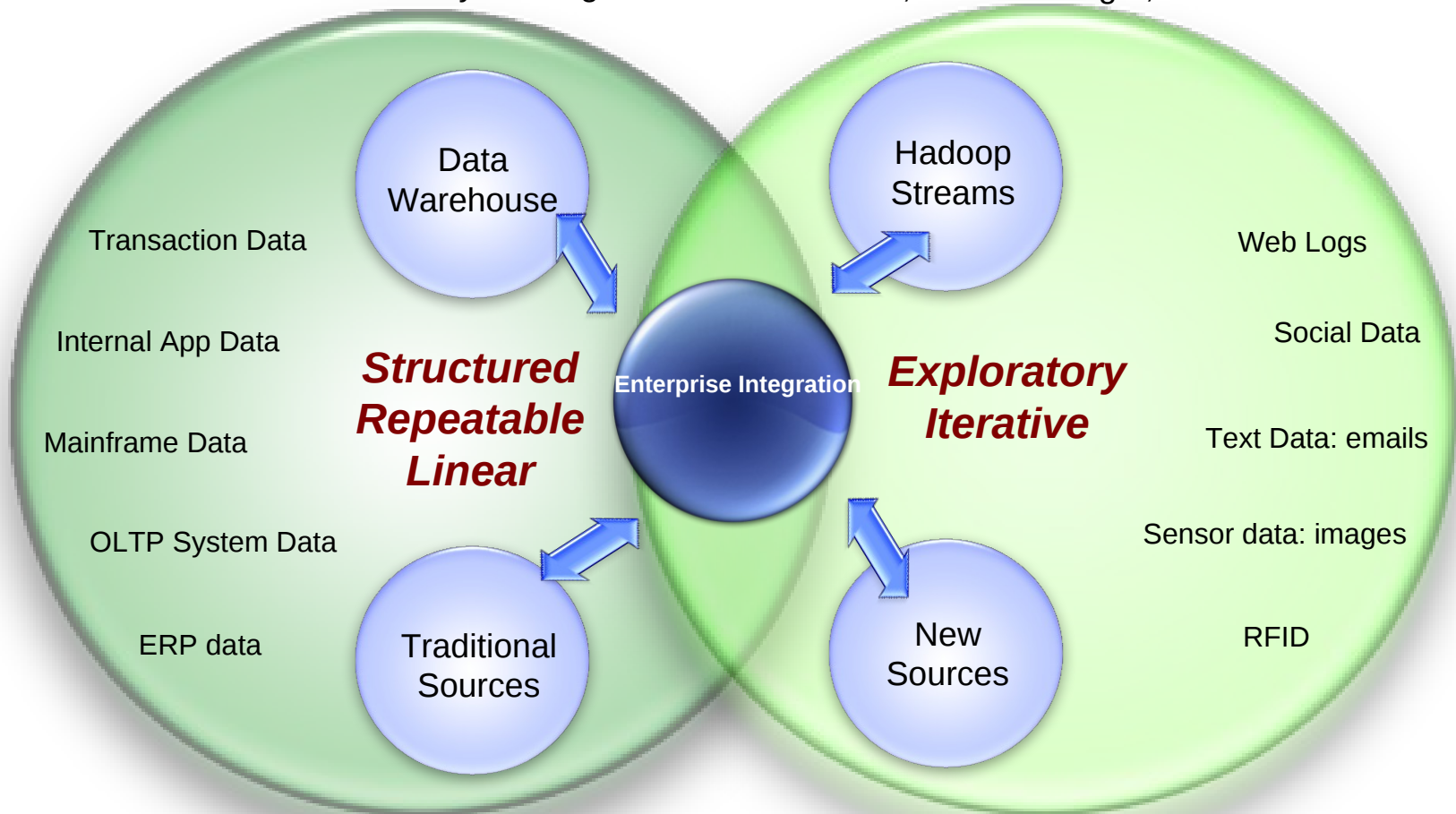
Integrating traditional with new approaches can unlock new insights

Traditional Approach

Structured, analytical, logical

New Approach

Creative, holistic thought, intuition



Integrating and Governing Big Data Challenges



Volume	Handling massive amount of data
Variety	Understanding the data – source, use, permissions
Velocity	Transforming and tracking the data
Veracity	Delivering trusted information at high speed
Vulnerability	Protecting Federated Data Sources
Vernacular	Internet Jurisdiction & Privacy Protection
Value	ROI Calculations without Product Descriptions

And without Mature Data Governance



Big Data will produce many high-speed mistakes

Because there isn't always one perfect outcome

The Predictive Governance Simulation

The Predictive Governance Simulation is a tool for evaluating the impact of Data Governance Policy on an organization in a safe test environment. We can change different policy options to evaluate both obvious effects and unintended consequences.

The Simulation was developed by the IBM Data Governance Council. Eventually, it will have 11 modules based on the 11 Categories of the Data Governance Maturity Model.

The first category available for testing is Data Quality, and all buttons lead to it so please test it today...



The Data Governance Council is creating Predictive Governance Simulations

While the Information Governance Community

Transitional Layer*		
Progress & Challenges	Business Value	
	Organizational Obstacles	
Levels	Name	Description
BD1	Initial	The initial phase is characterized by the application of existing Data Management methods and practices to Big Data technologies. IT Compliance Standards prevent adoption of all Big Data Solutions. ROI Control means each investigation must comply with structured business case methodologies. Data Stewards and Scientists are not part of this phase. Protection is also an afterthought.
BD1.1	Volume	Big Data is used in ETL pre-processing to prepare large volumes of existing data for faster ETL processing.
BD1.2	Velocity	
BD1.3	Variety	Organizations use existing varieties of data types and are very conservative about adding new streams without defined uses
BD1.4	Value	Organizations have structured models for creating business cases. Each investigation must have a defined product in mind with anticipated ROI.
BD1.5	Veracity	Since business cases are dependent on structured product development in initial stages of Big Data, Governance is an afterthought. Data sources are selected first based on perceived utility and veracity is not monitored or verified.
BD1.6	Vulnerability	Data Lineage is not performed and decisions are vulnerable to erroneous data. Applications are still the primary purpose of Big Data, so PII is not protected, and data sources or repositories are not encrypted either in transit or at rest. Data is highly vulnerable to manipulation or theft.
Transitional Layer*		
Progress & Challenges	Business Value	
	Organizational Obstacles	

Is Expanding the Maturity Model with a Big Data Category

IBM's Big Data Partner Eco-System – Over 70+



Big Data is Maturing Data Management

- **Big Data is a revolution**
 - Converging data and content worlds
 - Its a huge opportunity for all DM/ECM professionals
- **IBM is the leader in Big Data**
 - Tech Platform
 - Experience
 - Maturity Model
 - Predictive Governance

The screenshot displays the INFOGOVCOMMUNITY website interface. At the top, there is a navigation bar with the site name and a search bar. Below this, a 'Topics' section is visible, featuring three topic cards:

- Solutions:** A topic for people working with the Maturity Model, seeking insights and solutions. It includes a 'Make your submission' button and a 'View submissions & vote' button. The topic ends in 733 days.
- Community Advice:** A topic where practitioners can learn from others' experiences. It also includes submission and voting options. The topic ends in 733 days.
- Big Data Governance Maturity:** A topic asking if information governance includes big data and how it should be addressed. It includes submission and voting options. The topic ends 289 days until the final round.

On the right side of the page, there is a 'What's This?' section explaining the topic concept, and a 'Newest Submissions' section listing recent contributions such as 'Legal responsibility and data ownership' and 'A case for a Globally Represented Central Information Governance Council'.