



White Paper

Big Data Truths and Myths in the B2B Financial Industry

Many have acknowledged that big data is big and are willing to put resources behind it. We have observed numerous big data successes benefiting Business-to-Consumer (B2C) companies, but what about Business-to-Business (B2B) firms in the financial industry and financial technologies arena? We cordially invite you to complete the following True or False exercise in deciphering true values of big data in the right context of ‘fit for purpose’.

Please answer True (T) / False (F) / Unknown (?) for the following statements in the context of today’s B2B firms in financial industry or financial technologies arena.	Your Answer	See Our Answer at
1. Big data initiatives are for the primary benefit of improving client experience.		Page 2
2. B2B firms are innovators in adopting the latest and most advanced technologies.		Page 2
3. Transparency is much improved because B2B firms have lots of data.		Page 3
4. Veracity refers to perfect exactitude of data in one centralized data warehouse.		Page 4
5. Enterprise architecture projects are prioritized because I.T. knows the big data’s ROI.		Page 5
6. Tech vendors join forces with each other to seek higher values through big data.		Page 5
7. Big data emphasis on standardization and a single golden copy of security master.		Page 6
8. Data is counter intuitive, conventional wisdom always prevail in product development.		Page 6
9. There are well kept rosters for the # of data sources & how data is used and reused.		Page 7
10. Big Data supports tightening of risk appetite and favor more frequent risk reporting.		Page 8

Did you invite your technology, operations, product, finance, compliance, risk, and others teams (including your vendors and distributors) to jointly participate in this exercise? We ask as in all likelihood, a team effort will generally score higher due to a collective thought process rather than from an individual perspective. Calculate your total score using the following:

- Add 3 points for every correct answer
- Subtract 2 points for every incorrect answer
- Add 1 point for every question you answered ‘unknown’
(1 point is for your honesty and cautiousness. Keep in mind there are high sunk fixed costs in technologies investment, we help you discern truths and myths for your big data bet.)
- If you did invite others to jointly participate in this exercise, add 5 points to your score because you just overcome one of the biggest challenges in big data theology – ‘silos’.



Okay, let's compare our answers to yours beginning on the next page.

Warning: Big Data can tear our conventional wisdom apart and toss away our long-held intuitive thoughts. It opens up new opportunities and transforms the way B2B firms operate in the financial industry and financial technologies arena. You are about to expose to lots of ironies and contrasting ideas – that is how we come up with the better strategies.

Here the story begins ...

1. Big data initiatives are for the primary benefit of improving client experience. FALSE

The observed successes in the B2C arena are dominantly in the area of enhancing client experience; a direct transplant of the B2C approach to B2B does not necessarily yield the same benefits. One may argue that everything they do in business is about pleasing clients and it is a universal principal regardless of a firm being a B2C or B2B. Don't get us wrong as we are always a Champion in the Voice of Clients, but sadly the B2B reality suggests otherwise.

We all have seen B2B firms cut back on their service costs because it is not a revenue driver, nor new business development opportunity. Some adopt the mentality of "don't fix it if it is not broken". Heighten by the 2008 financial crisis, banks, broker-dealers, and insurance companies of all size are pressured to save costs wherever possible. They replaced high cost in doing relationship management the old fashion way with machines for automated call center solutions, shared services, and outsource many of their mid-and back-office operations. Are these really enhancing client experience or are they streamlining motives to save costs?

B2B are result oriented. You cannot lure a B2B firm to spend on unnecessary purchase because they like your warm and fuzzy service. B2B emphasis on value-adding and demand solid return on investment (ROI) for their projects. If big data cannot bring new products to sell, show new ways to beat competition, and cut costs or reduce risks, then what's the ROI in Big Data for B2B?

Enhancing client experience and better fact-based decision making should **not** be the end-goals or "primary" benefit of B2B's big data projects, but one of the means towards concrete business results.

To peruse at Data Boiler's ability to **crystalize** concepts for the **Big Picture** approach to **Big Data**, please visit our website at www.DataBoiler.com for use cases and details.

2. B2B firms are innovators in adopting the latest and most advanced technologies. FALSE

B2C retail banking may be the one constantly modernizing their channel distribution technologies from ATM, credit cards, on-line and mobile banking, etc. However the financial service industry in general tends to lag behind other industries for most advanced technologies because dependability always takes priority over innovations. Not only because there are high stake involved in financial transactions, but financial companies handle money or securities on behalf of others that potential misappropriation of clients' assets deem a significant risk.



Regarding B2B firms operate in convoluted capital market, retirement, insurance and corporate finance sectors the stake is even higher in dealing with institutional counterparties than the consumers. An unstable network in the B2B sector could cause huge loss to a company, or even have catastrophe effect on the overall market. Each day, B2B firms are combating with fragmented processes and deluge of data from variety sources. The complexity and high velocity of data (especially algorithm trading) in B2B also lead to veracity data problems. Financial engineers in the front office are creative in coming up with new products and/or new investment strategies, but middle-/ back-office are struggled with legacy systems that many are still running on mainframe computers using COBOL, FORTRAN, PL/1, Borland Delphi, Centura, CLIPPER, Clarion codes. B2B firms tend to be skeptical about cloud computing and open information exchanges. They are conservative in choosing technologies that are already stabilized and proven to be dependable. Thus we can conclude that B2B firms are **not** using the “latest” and “most advanced” technologies.

B2B firms interested in harnessing the power of Big Data may ask if they need to replace all of their office computers with tablets, latest mobile devices and buy expensive hardware. Upgrades should only be done as needed and the cost of Big Data platform is less than 1% of traditional database. It’s all about ideas instead of what expensive machine you own.

We won’t lure you to purchase expensive hardware. Check out our website at www.DataBoiler.com to see boiling hot ideas and use cases. Big Data can be as easy as slicing a piece of cake.

3. **Transparency is much improved because B2B firms have lots of data.**

FALSE

Things are definitely faster but **not** necessarily more transparent in the cyber era. Risk and Compliance control folks need to take root in deciphering the underlying and have slow-motion frame by frame pictures to see if there is any tolerance or non-conformance. We can understand how non-transparency may have occurred because of financial innovations, rather than firms trying to hide anything from stakeholders or intend to do anything wrong. We are able to help you discern truths versus myths.

Take for example the accumulator and convertor process to hold shares in omnibus account making the underlying position non-transparent. This process enables the available of no-load funds to investors. Just that the process to reduce number of accounts on mutual fund record system in rebating the broker/ dealers to offset client transaction fees also complicated the accounting, revenue management and fee billing process in the trade lifecycle.

Take another example, the ERISA 408(b)2 and 404(a)5 fee disclosure rules in the retirement space. It won’t be straight forward for plan participants to understand the various categories of fees and revenue sharing arrangements. More importantly, the rules call for accessing values of retirement services versus fees charged. A mere regurgitation of DOL Form 5500 reports won’t satisfy ERISA’s requirements, there ought to be appropriate benchmarking of investment performance when comparing fees for different investment options.



A third example is that there are different types of exchange traded funds (ETFs) beyond consider the style or investment category of the fund. An ETF can use full replication, optimization, synthetic swap and other representative sampling to benchmark a certain index. Also, exchange trade products (ETPs) can have different structures such as commodity pool, grantor trust, unit investment trust, mutual fund, or note. There may be security lending income for the ETF and possible affiliation with counterparties. All these represent different risks, tax exposure and other factors impacting the fund's performance. Thus not every ETP are created equal.

The list of non-transparency in financial market can go on and on, this added layers of administrative burden on the risk and compliance officers. While leveraging Big Data to do pre-cognitive detection of fraud could be the ideal, we have multiple ways to ease your **Risk and Compliance** administrative burdens. We are at your service to support your specific needs, please contact us to schedule an appointment to find out more.

4. Veracity refers to perfect exactitude of data in one centralized data warehouse. FALSE

Veracity is one of the four 'V's in Big Data according to IBM, it refers to "the level of reliability associated with certain types of data". The task is to "manage the reliability and predictability of inherently imprecise data types". Just that data reliability does **not** necessary mean the exactitude of data has to be "perfect", **nor** put everything into one "centralized" data warehouse.

According to [Big Data: A Revolution That Will Transform How We Live, Work, And Think](#) by Viktor Mayer-Schönberger and Kenneth Cukier, "looking at vastly more data also permits us to loosen up our desire for exactitude ... With big data, we'll often be satisfied with a sense of general direction ... we sacrificed the accuracy of each data point for breadth, and in return we received detail that we otherwise could not have seen ... we gave up exactitude for frequency, and in return we saw change that we otherwise would have missed ...".

This is an awakening call for us to give up devotion to exactitude. Haven't we all experienced typing not the exact words in Google, but the search engine is able to predict what we are looking for and auto-correct our typos? There's no sophisticated language algorithm behind Google in correcting our typos. The trick is simply because Google search engine has seen these misspelled words enough times, thus it is able to infer what people are usually looked for the right terms.

Nevertheless, big data is often "messy, varies in quality, and is distributed among countless servers" per Viktor and Kenneth. It does not make sense to centralize everything into one data warehouse. It's like putting all eggs in one basket. Also, by the time a common standard has been agreed upon, the business has changed. Therefore I.T. architecture ought to be agile for rapid responds to market changes.

Check out our website at www.DataBoiler.com to see how **I.T. Agility Optimization Services** can enable Big Data success with a lean and scalable architecture.



5. **Enterprise architecture projects are prioritized because I.T. knows the big data's ROI.** **FALSE**

We are **not** aware of any financial companies in B2B arena raised their earnings forecast because they are investing in Big Data initiatives. Interestingly there are market researches suggesting most of these firms are investing \$1 million to 10 million in average, or even \$100 million in forecast spending on Big Data projects. So, how is the ROI being measured for these multi-million dollar Big Data investments?

Improves enterprise architecture to better pull data are often cited as the justification to categorized projects as “Big Data”. We are not sure how much money actually goes into Big Data initiatives or overhaul legacy systems, but I.T. have worked industriously on various improvements that aren't always appreciate or understand by the business side. Anyway, one may have general impression that you have to have the chicken before the egg to investment in infrastructure before realizing any benefits from Big Data. We are not here to argue the egg or chicken first, but we disagree with false assumption that you have to pay big bucks in architecture before you can make any returns on your Big Data investments.

Resources deployment is one of the key aspects in strategic management, however many companies have leave their project prioritization process based on fierce fights among departments, top urgency from either seniors' or clients' requests. Not only does that counter-productive to the business, it creates unhealthy dynamics between dominant and sub-ordinate groups. Enterprise architecture projects “should” in theory be prioritized base on ROI, it is hardly the practice according to our observation.

We are a **passionate team** with **practical methods** to provide **people solutions**, please visit our website at www.DataBoiler.com to find our more. It is our commitment to enable your revenue growth, risk reduction, and cost containment, so you can realize excellent ROI through our engagement. This is what we mean by the true value of 212° - the extra degree to make a difference.

6. **Tech vendors join forces with each other to seek higher values through big data.** **FALSE**

Synergy is the gratifying reality of 1+1 = 3 or more through industrious collaboration of hard working teams. “If” tech vendors are industriously collaborating with each other, they would have openly shared their developments rather than emphasis on their proprietary technologies. It is evidence by their private cloud solutions that claims to be safer and better performance with APIs that are **not** interoperable with other platforms. We are not here to compare public and private cloud technologies, but we do know they are subjected to the same compliance and surveillance requirements per SEC.

The matter can be understood from a commercial point of view. Tech vendors try to protect their own turfs because there are high research and development costs in technology projects. Usually the payback period is very long for their investments. It would be hard for them to recoup money if their clients are too quick to adopt the latest open technologies, not to mention that if their developments are openly shared with others, it would jeopardize their bread and butter revenue stream.



In addition, financial technology conglomerates snapped up and acquired many emerging companies in the past few years. They took these emerging companies under their wings and often make their technologies more proprietary than system agnostic. Remember, tech vendors are “frenemy”; they fiercely compete and will only collaborate so long as their common clients have requested.

Check out our website at www.DataBoiler.com to see how **value chain management** services can align your vendors and use **infomediary / data aggregation** solution to seek higher values.

7. Big data emphasis on standardization and a single golden copy of security master. FALSE

According to Big Data: A Revolution That Will Transform How We Live, Work, And Think by Viktor Mayer-Schönberger and Kenneth Cukier, “big data is often messy, varies in quality, and is distributed among countless servers ...” This is the opposite to standardized data. Latest technology is capable of handling semi-structured, unstructured data or meta-data. It is able to predict what you are trying to look for in the search engine instead of requiring exactitude of input. Thus, it is **not** necessary to standardize everything.

Also, by the time a common standard has been agreed upon, the business has changed. I.T. architecture ought to be agile for rapid responds to business changes. We think it is not a problem for companies to have multiple sources of golden copies of security master. The matter is how companies can effectively and efficiently manage these multiple sources of data, rapid altering the source as it is applicable to the right situations, and properly control of the validity of data with audit trail. Enterprise service bus allows data caching, data persistence, data querying and messaging to be handled separately. It enables data aggregation from multiple sources, distributing data caching independent of OLTP, and performs OLAP dynamically for various report needs.

I.T. Agility Optimization services enable a firm to combat non-standardize technologies with much faster speed, more resilient platform, scalable developments and flexible choice of data sources. Please contact us to schedule an appointment to find out more.

8. Data is counter intuitive, conventional wisdom always prevail in product development. FALSE

In 2011, IBM's Watson computer beats human players on Jeopardy show. We aren't here to compare conventional wisdom with a machine, but the Watson story inspired us to re-think what is intuitive, rational or irrational. Dan Ariely, the New York Times best-selling author and professor at Duke, has used data to proven some irrational behaviors of human and how human intuitions can be wrong based on casualty. Data is objective and it is based on correlation. It can unveil hidden relationships and identify exceptions at areas where we human may miss because of oversight. It can crunch out certain answers quicker than average human. When the matter is beyond comprehension of a human, conventional wisdom would use gut feels to make an intuitive choice. Similarly, when matter is beyond the comprehension of computing algorithms, it can rely on a simple random or simulation to make “intuitive” choice. Thus, data is **not** necessarily counter intuitive.



Big Data can be both backward looking at trends, as well as forward looking to predict the future. Unlike statistical sampling, Big Data represents all or almost all of the people using a product. Big Data allows users to drill down data to individual level, hence able to substantiate and signal a market direction. As the data size increase, Big Data overcomes contextual problem with traditional statistical sampling. In addition, amazing things can happen by putting apparently unrelated data together (like the accidental invention of 3M Post-It), so why not let the data speaks for itself? We rather be humble and listen to the data than pretending to have God-like wisdom. The downfall of LTCM reminds us the danger to use causation approach to solve problem with complex algorithms and conventional wisdom. Therefore, we think conventional wisdom will **not always** prevail in product development, but complementing conventional wisdom with objective listening to Big Data can improve our human imperfection.

Our solutions include **Simulation / Data Modeling** and other techniques to harness the power of Big Data, please contact us to see how we can help you find new ways to beat competition, foster scalable development, and bring new products or value-added services to the market.

9. There are well kept rosters for the # of data sources & how data is used and reused. FALSE

A well-organized workplace will support a firm's risk and compliance controls. Every firm should have proper record retention policy, data and system inventory, clear roster to show number of data sources and how data is used and reused within privacy ordinance and other compliance rules and/or operational controls. Just that the workload of risk and compliance teams are overwhelmed by rapid growth of their firms through mergers and acquisitions before the crisis, the evolution of financial innovations that resulted in a convoluted market and fragmented processes, or heighten of regulatory requirements in post-crisis era.

We **doubt** the industry in general have well-kept rosters for number of data sources and detail document of how data is used or reused. Followings are our observations to support the argument: (1) Firms that are out of compliance with SEC 17a-4, FINRA 3110 and other rules are subjected to substantial regulatory fines. Unfortunately we have seen many cases become headline news in a bad way. (2) Firms are constantly combating with messy data problems every day. Managers from different departments do not trust each other's figures. Enhancement requests and/or incident reports are piling up to the roof and there also seems to be no end for the CTO to manage down numbers of spreadsheets. These are contradicting to the best practices of well-organized workplace. (3) We have seen business units have their own preference in contracting with data vendors and the lack of coordination in data procurement attributed to money wastage. It is a matter of accountability, everybody owns – nobody owns. It could become a vicious cycle when fire-fighting become the norm rather than the exception.

To rectify and improve the situation, please contact us to see how we can support your **risk and compliance** control processes. Also, you will benefit from out our **value chain management** services by rationalizing data sources, eliminating overlapped and/or contradictory data issues.



10. Big Data supports tightening of risk appetite and favor more frequent risk reporting. FALSE

In the post-crisis era, it is no doubt that regulators and firms are heightening their risk controls and requiring more frequent risk reporting. Just that it is **not** the direct results of Big Data **nor** Big Data support such tightening of risk appetite. Data in itself is un-bias and objective. It is human who interprets data choose to see thing as danger signal, or can choose to see the same thing as opportunity. Traditional statistical sampling failed to represent the entire populations could have create contextual problems that open up the analysis to debatable interpretation. In fact, Big Data overcomes such contextual challenge by inclusive of all or almost all of the data. By having full set of data, it enables the Big Data analysis to concurrently showcasing the mainstream direction, as well as niche opportunities with substantiation facts that can be drilled down to individual level.

A polling of market sentiment may signal economic recovery while there are still few dangers to be cautious, hence justifying the tightening of risk appetite and need more frequent risk reporting to closely monitor the market situation. In our opinions, we would say “thanks and no thanks to this kind of polling reports”. Thanks because it establishes the casualty for why we should be caution. No thanks because it doesn’t tell you where the danger or opportunities are, it doesn’t tell you what the correlations between events that leads to what consequences. In contrast, Big Data correlation may not tell you the casualty reason for “why” something is happening, but they alert us “what” and “where” things are happening. Hence, Big Data is able to complement our conventional wisdom to discern actionable insights and seek opportunities beyond the mainstream thoughts.

By the way, one of the key risks in market is credit risk. The word “credit” has its root from Latin – “*credere*” that actually means “to believe”. So rather than rejecting credit for casually reasons being generalized in the market, market participants should leverage Big Data to open their risk appetite in seeking “who” and “what” they can “make believe” of others (take educated risks) and revitalize the market with liquidity. Big data modeling provides justification to risk budgeting, expedite credit decision, and ensure healthy performance of loan portfolios.

Please visit our website at www.DataBoiler.com. Our suite of Big Data solutions and services includes:

- **Research and Consulting Services** to guide and manage clients’ Big Data projects
- **Simulation / Data Modeling Solutions** to support clients’ enduring business strategies
- **Datafication / Text Mining Solutions** to pluck diamonds from semi-structure or unstructured data
- **Infomediary / Data Aggregation Solutions** to cross-tabulate data for higher value purposes
- **Risk and Compliance Services** to ease the related reporting and administrative burden
- **Value Chain Management Services** to align vendors / distributors to work cohesively together
- **I.T. Agility Optimization Services** to enable Big Data success with a lean and scalable architecture



BIBLIOGRAPHY

Viktor Mayer-Schönberger and Kenneth Cukier, ***Big Data: A Revolution That Will Transform How We Live, Work, And Think***, Houghton Mifflin Harcourt, 2013

Paul Zikopoulos, Dirk Deroos, Krishnan Parasuraman, Thomas Deutsch, David Corrigan, and James Giles, ***Harness the Power of Big Data***, McGraw Hill, 2013

Jim Giles, ***Big data Lessons from the leaders***, The Economist Intelligence Unit Limited, 2012
http://www.managementthinking.eiu.com/sites/default/files/downloads/EIU_SAS_BigData_4.pdf

David Kiron, Renee Boucher Ferguson, and Pamela Kirk Prentice, ***From Value to Vision: Reimagining the Possible with Data Analytics***, MIT Sloan Management Review, SPRING 2013
<http://bf6837e6b8471f0219c2-9a1c5d9301ad213e0e491b46094b7800.r64.cf2.rackcdn.com/MITSMR-SAS-Data-Analytics-Report-2013.pdf>

Erik Brynjolfsson, Andrew McAfee, Paul Barth, Randy Bean, Michael Schrage, Jim Stikeleather, Gary Hawkins, Alex (Sandy) Pentland, Chris Young, Niko Karvounis, Irfan Kamal, Simeon Vosen, Torsten Schmidt, Tom Davenport, Thomas C Redman, Michael Schrage, Dominic Barton, David Court, Bill Franks, Dennis Crowley, Justin Fox, Robert Plant, Rick Reynolds, Robert Griffin, and Dana Rousmaniere, ***The Promise and Challenge of Big Data***, Harvard Business Publishing, 2012
http://docs.media.bitpipe.com/io_10x/io_109281/item_671761/The%20Promise%20and%20Challenge%20of%20Big%20Data.pdf

Paul Barth and Thomas H Davenport, ***Big Data Executive Survey Themes & Trends***, New Vantage Partners LLC, 2012
<http://newvantage.com/wp-content/uploads/2012/12/NVP-Big-Data-Survey-Themes-Trends.pdf>

Nathaniel Rowe, ***Big Data for Small Budgets***, Aberdeen Group, December 2012
http://info.talend.com/rs/talend/images/WP_EN_BD_BigDataSmallBudget.pdf

Jeffy Kelly, David Floyer, Dave Vellante and Stu Miniman, ***Big Data Vendor Revenue and Market Forecast 2012-2017***, Wikibon, February 19, 2013 (updated April 17, 2013)
http://wikibon.org/wiki/v/Big_Data_Vendor_Revenue_and_Market_Forecast_2012-2017