Shifts & Twists in Business Analytics:
Reflections from Qlik Qonnections and Alteryx Inspire

Transfăgărășan Pass in Romania by Cristian Bortes, Flickr #1414351112—CC BY-NC-SA 2.0

Like driving up a mountain pass, Business Analytics evolves mostly upward, but with unexpected shifts in thinking and sudden twists in technology. In this article, two recent events provide clues about the shifts & twists, plus when we might summit this mountain pass. An echo from my young son rings in my mind, "Daddy, are we there yet?"

> **Abstract**: At recent vendor briefings, I observed seven industry changes with analytic systems that may indicate fundamental shifts in thinking and twists in technology over coming years. Given are suggestions for how organizations can be ready for these changes.

As a technology analyst covering the Business Intelligence (BI) industry for several decades, I have witnessed many changes in Information Technology (IT) in general and in Business Analytics (BA) in particular.

In early days, the state-of-the-art analytic technology consisted of financial report generation with a few counts and totals. In other words, it was **describing known data**, which has been the driving force behind the BI industry with its real-time performance dashboards and interactive visual analysis tools.

More recently, analytic technology is now driven by the rapid evolution of machine learning algorithms, especially those based on artificial neural networks. More importantly, emphasis has shifted toward **generalizing beyond known data**. Those changes have been fascinating and often unexpected, as they are shifting and twisting through the BI landscape.

Touting their newest products, the decades have bought a constant parade of BI vendors, much like a holiday parade, with horns blaring and fortunes soaring. Cheers arise from the crowds with each new technology wave, promising to solve our business problems.

Qlik
Qonnections in May 2019—Alteryx Inspire in June 2019

Recently I attended analyst briefings by Qlik and Alteryx to hear their latest & greatest, expecting to witness more parades.

Another Shift & Twist?

However, that was not what happened! The Qlik conference highlighted the acquisition of Attunity **the week before**. [01] Likewise, the Alteryx conference was greeted, **on its first day**, with the acquisition of Tableau by Salesforce. [02] These parades were unusual…

Qlik acquired Attunity for over $600M from private capital. I wondered… Why did a dashboard company acquire a company for its enterprise-level data pipes? Just because they can? You can only put so many pixels into those dashboards.

The conference buzz swirled around those large acquisitions, with some implying that the BI industry is now *for-adults-only*. Was BI innovations over for technology-eccentric entrepreneurs and à la cart vendors? Is the ante now over a billion to play in the BI game?

On the surface, these trends appear as the same-old tech remixed and consolidated, amid the confusion of bigdata-galore, clouds-everywhere, insights-mania and analytics-aplenty. Are these developments significant and fundamental enough to hint of a new shift or twist?

What are my key observations at these two events? The sections below expand on seven observations, plus suggesting practical Take-Aways.

- · **Driver**—Analytics is evolving rapidly but is not evenly distributed
- · **Barrier** —Analytics requires a solid DataOps infrastructure
- · **Barrier**—Analytics requires new thinking but most lack of basic literacy
- · **Multiplier**—Analytics should mature isolated apps to integrated systems
- · **Decision**—BI vendors face critical decisions about analytics
- · **Enabler**—Business/data analyst must play enabling roles with analytics
- · **Energy**—Fuel to power the above is coming from next-gen data geeks

The sections below expand on these observations, plus offering Take-Away suggestions for managers.

Driver—Analytics Is Rapidly Evolving and Is Not Evenly Distributed

The dominant theme of both events was analytics. So, I asked everyone whether, how and why they were using analytics. I lost count of the number of times persons used the analogy of teenage sex to describe the use of analytics in their organizations. *Everyone is talking about it, but…*

My assessment is that the majority are using conventional machine learning (like random forests) for decision-support projects with curated single-use datasets, often lead by outside experts. However, **state-of-the-art analytics currently has rapidly evolved beyond conventional machine learning techniques to be almost a separate category**. Only 3–5% of *normal organizations* [03] are using the current generation of analytics and are ready to leverage the next generation of analytics.

The term that I use for these future analytics is next-gen analytics that is based on the Lego-like zoo of neural network architectures, consuming large data streams and requiring considerable tensor processing power. One can think of resulting models as **information sponges that are able to distill many gigabytes of raw data into a million weights** (simple numbers). These weights can then be used for a variety of tasks, such as making a single prediction of future sales, categorizing images for cancer detection, translating Russian speech into English text, perform a complex sequence of robotic motions, or generating a sample of the initial dataset.

Unfortunately it is rare for a company to have the enterprise-level capability to fork pre-trained models from github, customized the code in Jupyter notebooks, retrained via transfer learning with their unique datasets, deployed into production as a web service, integrate with existing IT infrastructure, realize actual business value, and govern with confidence. This is not your Grandma's supervised learning anymore. Yet **all the essential tools and skills are readily available today**, for any curious techie. [04]

So, why are these future analytics not evenly distributed?

Barrier—Lack of A Solid DataOps Infrastructure

As mentioned, analytic examples that I observed at these events were usually self-contained analytic applications with a manually curated single-use training dataset. Further, **data prep typically took a month, while data analysis took days**, which implies…

The reason that most companies are only talking about analytics is that **they struggle with basic data capture and integration**. It seems ironic that our IT industry has spent seventy years struggling with the general data integration problem, making great strides but never quite reaching the finishing line.

Organizations that embraced the lessons of Data Warehousing in the 1990s are now the ones who are doing analytics today. **The insatiable appetite for data as examples of business situations to fuel learning analytics has been a surprise to all and a blessing to those who invested early into DataOps.** [05].

Without a solid DataOps infrastructure, each analytic project is a start-from-scratch effort to create the labeled train/test dataset. This should be performed in a single SQL statement. When successful, deployment creates yet-another independent stovepipe system, floating somewhere in the cloud. This should be part of normal production workloads administered centrally.

Why did Qlik acquire Attunity? …because of their well-honed enterprise-level DataOps. Why did SalesForce acquire Tableau? …because SalesForce had lots of data and solid DataOps infrastructure but needs a better way to consume analytics. Why is Alteryx well positioned for analytic systems? …because they have user-friendly DIY DataOps tool, can easily add analytic modules (like yet another ETL transform) to their toolbox, and deploy a data-pipe-to-web-service in a single click.

Take-Away: An insightful exercise for an executive to determine if their company is ready for analytic systems is the following: Ask for a simple logistic regression analysis on recent manufacturing defects (or similar). What is the root cause of these defects? Ignore the results but note how long it took. Ask specific questions about duration, person-hours, and steps required for this analysis, especially data prep. Then, appear excited and request deploying this analytic model into production ASAP. Be prepare for someone to wet their clothes. Finally, multiply by x10-x100 to estimate resources and duration for a realistic analytic system. In an analytic-mature company, this estimate should be one week with a couple of persons.

Barrier—AI Literacy for Everyone About Data and its Analytics

I was thoroughly impressed with the effort that Qlik has made in the open Data Literacy Initiative. [06] They expended the resources and talent necessary to do it properly and share with all. I would urge EVERY organization in the BI community to support this data literacy initiative (or something similar).

The next step is to extend basic data literacy with a heavy dose of analytic literacy. **Hiring a few talented data scientists will not sufficiently change organization culture to ensure success for your development team**. Everyone should be thinking and discussing issues with common terminology and concepts.

There are several subtle shifts with analytic literacy. Here are two key shifts. For details, see [07]

First, essential is understanding the shift from *describing known data* to *generalizing beyond known data*. This is the essence of being analytic! It is a **conceptual leap that everyone performs naturally and quickly, but also predictably irrational and sub-optimal**. [08]

Second, there is the shift is *crafted (or static) logic* (which is the norm for decades) to *learning logic*. In the latter, applications are created from analytic models that are trained and validated (called *Fit*) with a training dataset containing labeled (or categorized) examples. Then used to infer an output from an input (called *Predict*). This Fit-Predict dance is called *Supervised Learning* (although it is really training, not learning).

Take-away: A first step is to define your terminology about analytics and related topics. It is changing monthly. Start with *Fit* and *Predict*, and continue with other concepts critical to your analytic initiatives. Ensure that everyone is speaking the same language.

Take-Away: Finally, ensure that your corporate messaging is in sync with the above. Especially, define your specific usage of the term AI or avoid it entirely. [09]

Decision—Strategic Options for BI Vendors re Analytics

Wandering through the exhibits, I queried vendors about how they were supporting analytics for their customers and how they were embedding analytics within their products. The signage for every vendor stated something about analytics; however, only about half were supporting (or embedding) canned machine learning modules, such as logistic binary classifiers, K-means clustering, and various decision trees. There was only one vendor (a global system integrator) who spoke intelligently about neural networks and understand their significance and differences. When asked how many their client projects were applying neural technology, he responded "*none*"!

This presents a dilemma for BI vendors generally. Their strategic options concerning next-gen analytics are:

1. **Ignore**—No revenue in short term. Too risky for early adoption. Business is currently doing great.
2. **Hype**—It's the next big tech wave. Let's bet our company's future on it. Label all current products and services as *AI*.
3. **Lead**—Be a thought-leader. Openly recognize potential and limitations. Educate everyone. Think readiness.

As a BI vendor, which option is the best one for you?

I recommend option #3—**Lead**, but with applying prudent business savvy. As a thought-leader, put resources into increasing the literacy of data and of analytics internally within your organization, externally with your customers, and generally within your industry. Identify and maybe eliminate the key barriers to adoption of next-gen analytics within your organization. Go slow on big bets, but try many tiny bets to learn best practices for your organization.

Take-Away: Initiate a small analytic innovation team to scope the problem by: (a) educating themselves, (b) compiling valid use cases, (c) implementing 2–3 prototypes, and (d) creating a readiness plan for executive discussion. Establish a six-month plan with clear monthly milestones.

Take-Away: Start monitoring analytic research stream as centered on arXiv.org because there is so much posted weekly. Also, the latency from research to application is steadily declining.

Assign a talented person who can scan academic research and extract practical insights. This person should spend at least 20%-time monitoring, distilling, and regularly briefing. To assist, open aggregating services are steadily improving, such as PapersWithCode, arXiv-Sanity, Two-Minute Papers, O'Reilly AI Newsletter, and Semantic Scholar. [10]

Take-Away: Being a thought-leader also implies being professional responsible for the growing social implications and ethical issues of analytic systems. Otherwise, this is similar to the negligence of theoretical physicists in 1940s who were evangelizing nuclear power generation while ignoring dangers of weaponizing nuclear technology. Time to seize this leadership opportunity! Make the ethics of analytics an acceptable professional topic to discuss within your organization. Collaborate with other organizations, and pool the wisdom about AI futures. [11]

Multiplier—From Developing Applications to Deploying Systems

As mentioned earlier, most analytic examples that I observed were self-contained decision-support applications trained with a single-use curated dataset.

To reap the greater business value from analytics, attention should be redirected **from developing applications for a single analytic use case to deploying systems enhanced with multiple analytic modules**. A conspicuous analytic application attracts attention, but insignificant analytic modules embedded within enterprise systems can be major multiplier of value.

Think of these analytic modules like ETL modules from Informatica in the 1990s or ELT transforms from Teradata today. They are customizable generic transform functions mapping A to B, where the A and B are now wildly different. If properly implemented, the system modules will incrementally enhance organization behavior toward customers, suppliers, and others.

For over a decade, BI has steadily moved toward the data consumer being the data developer, learning the technical skills necessary and redesigning tools to be easy-to-use. The term *Pervasive BI* is often used to denote this trend. Vendors have catered with various forms of self-service BI. The latest is assisted analytics, where the analytic tool suggests appropriate charts to visualize based on the nature of the data.

In general, assisted analytics is a healthy trend in maturing BI tools. The problem is that assisted analytics distracts from the larger issues by focusing on the small issues within model development. You see the trees but lose sight of the forest. The larger issues (the forest) deal with the business use case and its analytic value proposition, along with the specific technical use case, its parameters, and desired metrics. More rigor and validating should be embedded in the tools or even methodologies for supporting analytic systems work.

Take-Away: Upgrade your DevOps and DataOps agile development methodologies for future analytic systems. This should be an industry-wide effort. Once a standard emerges, vendors can create the tools required to support this methodology. If this analytic development methodology is constructed properly, deployment to production should be a single click (plus copious testing). Lean on your analytic vendor partners to assist.

Enabler—Role of Business/Data Analyst is Critical

At Qonnections, I was asked to do a short introduction explaining AI, prior to a talk on *Augmented Intelligence by Qlik Sense* by Elif Tutuk and Vinay Kapoor. I struggled about what I could say substantively in ten minutes. I thought about the nature of the audience, which was mainly data analysts (tech-savvy) with a few LOB business analysts (biz-savvy). The answer then seemed obvious. There was a very important message to deliver!

There is a **growing cultural disconnect between IT professionals (especially business analysts) and data scientists**. They speak different languages. Unless resolved, the outcome will likely hurt the organization by limiting its BI evolution. **To do their jobs, data scientists will wander off and implement stovepipe systems, floating in the clouds and free from IT involvement**. And, **business analysts will disown any responsibility or knowledge about those AI systems**. Data analysts will continue to dabble in visual charts describing historical data. **Executives will be frustrated with managing two IT groups who do not communicate and with maintaining incompatible systems**. And, no one will perform the hard work of conceptualizing the business use cases for analytic modules. Everyone will be unhappy!

Over the history of BI in the enterprise, strong executive leadership is most needed now to navigate the entire organization through this next period of BI evolution. **The balance between governance and agility has never been more delicate. Change in IT roles and responsibilities never more dramatic. And, the rapidity of technology change has never been greater.**

The data science team must focus on the bits, ensuring that analytic models are properly architected, trained, and validated again and again, amid rapid advancements in neural network technology. They need clear statements of use case objectives and criteria, along continuous model evaluation after deployment. Data scientists are the wrong persons to craft those requirement statements, which require an intimate understanding of the business. It is management, from top to middle, who are responsible for those statements.

Finally, guess who should be bridge between all levels of management and the technical specialties (especially data science), ensuring that the objectives and criteria for analytic use cases are accomplished properly.

The point is… **The business/data analyst is the enabler for successful analytic systems of the future**. [12]

Take-Away: Do a strategic reassessment of the role of business/data analyst across the enterprise. How are these people currently being used? What is the balance of biz-savvy with tech-savvy, relative to their organizational position? How should their role evolve to support future analytic systems? What new responsibilities should they have? And, what new skills and tools do they require for these new analytic-centric tasks?

Take-Away: Designing a program to train the next-gen business/data analyst to upgrade their skills and aspirations for this emerging new role with analytic systems. If successful, they should become best buddies with the data scientists, rather than trying to substitute for them as citizen data scientists. [13]

Take-Away: Think strategically about the function of the data science team. Do you really want to hire an entire team as employees? Maybe a couple coordinating data scientists would be sufficient. Custom analytic work could be out-sourced to highly specialized development groups, while common analytic tasks could be web-serviced from cloud vendors. If so, there will be a feeding frenzy by the global SI firms, plus AWS and Azure.

Energy—Curiosity from Next-Gen Data Geeks

I immediately observed that attendees at both events were curious, energetic, motivated, and under 40. It was a bit jarring for me to enter the room and dramatically increase the average age. There was revival energy in the attentiveness of the audience, like this technical stuff really mattered to them personally. But, why?

Over breakfast, I talked with a random Alteryx customer. He said without prompting, "*Using this product is fun and engaging. I really like the artistic details in the icons. Makes them distinctive. Easy to learn. Pleasant. My workday goes fast*."

His mood echoes the energy from SAS User Groups love fests in past decades and Tableau rock concerts of past years. The energy seemed to be coming from their gut, not their mind or wallet. Here is my explanation to the WHY behind this youthful energy…

First, **their motivation has shifted from tech-centric fondness to story-centric curiosity.** They do not care about the tech, just about acquiring the tools and skills to tell stories based on the data. They do not care about analytic-generated insights by themselves, but about the stories that those insights foretell.

Second, **analytics adds *color* (more depth and drama) to their stories**. Analytics is shifting attention away from insights into "what is interesting or significant" toward knowing about who caused what and why. It is like a *who-done-it* mystery that investigates how Ms Helen was killed. Their role is becoming like that of an investigative reporter who has a 6pm news deadline.

Third, **any story does not end with *who-done-it*. There must be resolution.** My favorite question is, "*What do you do with what you know?*" In other words, the story must end with the specific actions that should occur to resolve the business problem. This will require knowing the relevant casual relationships to be employed. [14] Or stated simply, the next-gen data geeks want to understand how the business works in order to finish their story.

Take-Away: Executives, support your next-gen data geeks by validating their contributions to the organization. Too often, their contributions get buried several levels down or never come to the attention of management.

Take-Away: Next-gen data geeks are very aware of the social implications and ethics issues of their job responsibilities. They realize that it is now professionally negligent to ignore or be naïve about these issues. It is now considered part of their story-telling. Executives should support this trend by promoting open and honest discussions about these issues and making these discussions a safe and normal part of the organization culture.

Are We There Yet?

My opinion is that we are nowhere close to the mountain pass. For several more decades, the BI evolution will continue upward with its surprising shifts and twists. So, tighten your seat belt. Enjoy the views!

Here are a few final suggestions that summarize the previous…

Take-Away: Once you have a viable analytic use case, shift your thinking from analytic-centric applications to analytic-driven systems. Because of the value multiplier, analytic use cases that are properly implemented will start to change organization behavior externally. A conspicuous analytic application is nice, but several insignificant analytic modules embedded within the enterprise system can have huge business impacts. Deal with the issues of deployment, operations, and governance.

Take-Away: Forget what you think you know about analytics. Analytics is evolving daily, and the latency from applied research to practical applications is rapidly declining. For instance, research papers are posted on arXiv in the morning; successful github forks are reported in the afternoon, and web service available the following day. Potential use cases for off-the-github next-gen analytics now exceed the ability of industry to absorb these new BI innovations.

Take-Away: To expand the thinking of your team about next-gen analytics, try this exercise: Click here ;)

Attributions

I received many useful suggestions on the article draft from the following persons: Remco Broekmans, Rob Gerritsen, Dan Graham, April Healy, Julie Hunt, Dave Imhoff, Todd Margolis, Joseph Paolantonio, Neil Raden, and Elif Tutuk. Thank you for helping me untwist my thoughts!

Notes

[01] Qlik press release at https://www.qlik.com/us/company/press-room/press-releases/qlik-to-acquire-attunity. Plus an article by Wayne Eckerson of Eckerson Group at https://www.eckerson.com/articles/qlik-acquires-attunity-fleshes-out-its-data-analytics-platform. Eckerson states that Qlik "wants to be a one-stop shop for integrated data and analytics for large enterprises".

[02] Salesforce press release at https://investor.salesforce.com/press-releases/press-release-details/2019/Salesforce-Signs-Definitive-Agreement-to-Acquire-Tableau/default.aspx. Plus an article by Forbes at https://www.forbes.com/sites/patrickmoorhead/2019/06/18/salesforces-tableau-acquisition-admitting-organic-innovation-failure/#1e26457b3ca2.

[03] Controversial book entitled “The Big Nine” by Amy Webb (March 5, 2019) that states, “The big nine corporations—Amazon, Google, Facebook, Tencent, Baidu, Alibaba, Microsoft, IBM and Apple—are the new gods of AI and are short-changing our futures to reap immediate financial gain.” Useful (but debatable) insights into the technology, economics, and politics of the current AI hysteria, along with well-documented facts. https://www.amazon.com/dp/B07H7G7CMN/

[04] Forthcoming article on next-gen analytics. Will post a link here when published in Towards Data Science. Anyone interested in collaborating?

[05] Wikipedia entry has useful background on DataOps at https://en.wikipedia.org/wiki/DataOps. However, search for more current and detail material elsewhere, such as the Eckerson Group Thought Leadership section at https://www.eckerson.com/thought-leadership.

[06] Data Literacy Initiative at https://thedataliteracyproject.org/. Browse the site. Take the online assessment.

[07] Here are several resources about analytic literacy:

- https://towardsdatascience.com/data-science-literacy-for-the-enterprise-fadaf9268494
- https://www.gartner.com/smarterwithgartner/a-data-and-analytics-leaders-guide-to-data-literacy/
- https://tdwi.org/events/onsite-education/onsite/sessions/business-analytics/adv-all-developing-data-analytics-literacy-workshop.aspx
- https://hbr.org/2018/10/your-data-literacy-depends-on-understanding-the-types-of-data-and-how-theyre-captured

[08] Delightful classic book entitled “Predictably Irrational” published in 2008, found at https://www.amazon.com/dp/B002C949KE/. Worthy of a read every few years, plus search for books and articles that refer to this classic.

[09] I hate what the term AI has become to mean. For more of my ranting, check out “Vendors—Define Your Usage of #AI” at https://towardsdatascience.com/vendors-define-your-usage-of-ai-9495b30ebd28

[10] The full URL for the open arXiv aggregating services are:

- https://arxiv.org/
- https://paperswithcode.com/
- http://www.arxiv-sanity.com/
- https://www.patreon.com/TwoMinutePapers
- https://www.oreilly.com/ai/newsletter.html
- https://www.semanticscholar.org/search?q=deep%20learning&sort=relevance&fos=computer-science

[11] AI Ethics Resources are:

- Partnership with AI at https://www.partnershiponai.org/
- fast.ai AI Ethics Resources at https://www.fast.ai/2018/09/24/ai-ethics-resources/
- Allen Institute for AI at https://allenai.org/index.html
- Machine Intelligence Research Insitute (MIRI) at https://intelligence.org/
- World Economic Forum on Strategic Intelligence at https://intelligence.weforum.org/

[12] Forthcoming article on the new roles for the business/data analyst. Will post a link here when published in Towards Data Science. Anyone interested in collaborating?

[13] Be careful of *Citizen Data Scientist* initiatives. Only a few next-gen data geeks will want to tweak the bits of tensors, but not the majority. This is like retooling race car drivers into automotive engineers. The organization needs both, working in close partnership. [13]

[14] Forthcoming article on the emergence of causal analytics. Will post a link here when published in Towards Data Science. Need to research literature by Judea Pearl (https://en.wikipedia.org/wiki/Judea_Pearl) and others. Anyone interested in collaborating?