Production Analytic Platform—Into the Future

JANUARY 2018

A ThoughtPoint by
Dr. Barry Devlin, 9sight Consulting
barry@9sight.com

As might be expected of any architectural thinking, the Production Analytic Platform must look beyond current issues and solutions to how business and technology environments may evolve. The final ThoughtPoint in this series considers how advances in analytics and artificial intelligence influence this approach.

A thin line between analytics and AI

The future of analytics is closely aligned to artificial intelligence. In fact, given that both are simply ways of building and exercising models of the real world, the boundary between them is paper thin and growing more porous daily.

Thirty years ago, the first Data Warehouse architecture[1] emerged from the convergence of business demands for improved access to consistent management information and the new possibilities offered by relational technology. With it, BI was born and has become the lens through which decision making is still viewed. Some ten years ago, advances in business needs and technology combined to shift our focus again—to big data, predictive analytics and Hadoop.

Now, even as the foundations for fully operational analytics are being laid with the Production Analytic Platform[2], technology is undergoing yet another tectonic shift: Artificial Intelligence (AI) is supplementing and, in some case, supplanting BI and analytics at the heart of decision-making support.

AI—and an array of partial and overlapping synonyms[3] such as *machine* and *deep learning*, *cognitive computing* and *neural networks*—is a very broad and often confusing topic, ranging from theoretical mathematics, through neurobiology, to fundamental computer science. For anyone with a BI/analytics background, one simple description is as technology that enables machines to learn about and model the real world, to make decisions about it, and act accordingly, minimizing the level of human intervention in the entire process. Autonomous vehicles provide an obvious and easily understood example of both independent learning and automated decision making / action taking.

SPONSORED BY
www.teradata.com TERADATA.

This explanation emphasizes the thin line between analytics and AI. There are, of course, differences in techniques and tools, for example in the use of neural networks and platforms such as TensorFlow and OpenAI. However, the intent of the AI process is, in fact, an extension of analytics: From searching for the "right" answer among trends to discovering intelligent options and ranked possible solutions to real world challenges. The continuity of this spectrum of function can be seen, for example, in the range of function available in the Teradata Analytics Platform— from SQL to advanced analytic techniques like machine learning, graph, pattern, pathing, and sentiment, to full AI tools such as TensorFlow and Spark on the near-term horizon.

AI and the Production Analytic Platform

> As the Production Analytic Platform evolves, AI will become more important, complementing current operational and analytic aspects. Indeed, the evolution of AI will, in part, drive the future of the Production Analytic Platform.

With such a thin line between analytics and AI, it is immediately clear that the Production Analytic Platform must and will include elements of AI libraries and platforms. Even a high-level review of AI tools[4], shows significant overlaps between traditional mining and new deep/machine learning function in many available platforms and libraries.

The emerging Production Analytic Platform therefore spans the entire spectrum of IT processing—from operational systems, through BI and analytics, all the way to AI. It brings together data and function from the initial collection and cleansing of data and information from every environment, internal and external, through simple reporting and problem solving, to deep analytics, both predictive and prescriptive. It this way, the inclusion of AI in the Production Analytic Platform helps close the MEDA loop described in part three of this series[5].

As AI-supported function becomes more common, analytics and basic AI become mere table stakes in business competitiveness. New business needs emerge, driving further advances in technology that enable enterprises to differentiate. This symbiotic relationship between business and IT is the ***biz-tech ecosystem*** described in *Business unIntelligence*[6] that is the foundation of digital transformation.

Future AI also drives the Production Analytic Platform

The significant advances seen in AI in the current decade—particularly in image recognition and natural language processing—rest on two key foundations: (i) the exponential growth in processing power and data storage of modern hardware and (ii) the even faster growth of data/information training sets available from the Internet. Together, these trends have led to significant centralization of AI processing with the problems entailed in the transport of such data from the edges of the network to the center.

Recent developments in AI are focusing on reducing both the training data required and processing power needed. DeepMind's AlphaGo Zero game, for example, learned to play without *any* training data and uses considerably less power than its predecessor[7]. While still in the research lab, the direction for business AI software is clear.

In addition, vendors are increasingly pushing AI function—both training and operation—to the edge of the network. With localized function in smartphones and autonomous vehicles, business users experience enhanced decision-making timeliness and IT can significantly decrease data transportation costs and delays.

It follows, therefore, that these complementary business and IT needs will drive the Production Analytic Platform toward an increasingly decentralized and distributed foundation. This direction is already visible in the Teradata Everywhere™ approach which provides the full power of the Teradata Analytics Platform and open source software across an array of deployment options, from on-premises to the edge of the public cloud, built on Teradata or commodity hardware.

Conclusion

Decision making and action taking today and into the future demand tight integration of very different types of data and function, from collection and analysis, through predicting future states, to taking immediate, operational action. The Production Analytic Platform, based on relational database technology, offers the most appropriate and effective solution.

As a modern logical architecture, the Production Analytic Platform addresses significant changes in data usage patterns and technological advances over the past half-decade. Further change in both business and IT is expected and the platform will evolve to meet this ongoing transformation. The key aspects of this architecture are:

- *Relational database technology* at the core—for reliability, availability, scalability, maintainability, and performance levels spanning operational and analytical usage
- *Built-in non-relational support* for storage of data in a wide range of formats (e.g. JSON, Avro, etc.) and processing approaches from simple query, through analytics and AI, accessible via SQL as well as native languages by users of all types and skill levels using their preferred tools
- *Data virtualization support* for access to and use of data stored remotely in both relational and non-relational formats, optimized by caching and other means
- *Distributed and decentralized operation* to support the longer-term migration of AI to the edge of the network with local data and processing

The four characteristics above are not new. Today, however, they often exist across different products and separate platforms, leaving enterprises to employ or adopt the role of systems integrators. The Production Analytic Platform integrates these functions in a single platform.

Users of analytics and AI are becoming ever more varied in their goals, use more data of different types and sources, and demand faster responses to both operational and informational processes. Indeed, individual users increasingly require seamless transmission of information and expertise from developer to data scientist to business analyst, and *vice versa*. Operationalizing experimental discovery work demands simple and elegant transitions between these roles. Existing IT environments—a mix of operational systems,

data warehouses / marts, and data lakes that have grown historically more complex in management and use—are unable to meet these expanding business demands.

Adding yet another technology environment to this mix is not an answer. Rather, a more realistic and achievable approach is to choose an existing environment and expand and improve its functionality to address these increasingly central business needs. The wide-ranging strengths and maturity of the relational database—especially as it has evolved to support enterprise data warehousing—offer the best starting point for this journey. The Production Analytic Platform defines and describes the starting point and the target, both in the immediate future and in the longer term. Better still, the technology already exists to start that journey today.

Dr. Barry Devlin is among the foremost authorities on business insight and one of the founders of data warehousing, having published the first architectural paper on the topic in 1988. With over 30 years of IT experience, including 20 years with IBM as a Distinguished Engineer, he is a widely respected analyst, consultant, lecturer and author of the seminal book, "Data Warehouse—from Architecture to Implementation" and numerous White Papers. His book, ***"Business unIntelligence—Insight and Innovation Beyond Analytics and Big Data"*** *(http://bit.ly/BunI-TP2) was published in October 2013.*

Barry is founder and principal of 9sight Consulting. He specializes in the human, organizational and technological implications of deep business insight solutions combining all aspects of internally and externally sourced information, analytics, and artificial intelligence. A regular contributor to Twitter (@BarryDevlin), TDWI, BACollaborative, and more, Barry is based in Cape Town, South Africa and operates worldwide.

[1] Devlin, B.A. and Murphy, P.T, *"An architecture for a business and information system"*, February 1988, IBM System Journal, 27(1), http://bit.ly/2hfgiVl

[2] Devlin, B., *"And now… a Production Analytic Platform"*, October 2017, http://bit.ly/2k3QmQx

[3] Vorhies, W. *"Machine Learning—Can We Please Just Agree What This Means"*, December 2017, http://bit.ly/2yh11J3

[4] Davis, S. and Baker, J., "Top 12 AI Tools, Libraries, and Platforms", September 2017, DZone, http://bit.ly/2AwQgbS

[5] Devlin, B., *"Production Analytic Platform: A Shrinking Decision Cycle"*, December 2017, http://bit.ly/2AHSgsE

[6] Devlin, B., *"Business unIntelligence"*, (2013), Technics Publications, New Jersey, http://bit.ly/BunI-TP2

[7] *"AlphaGo Zero: Learning from scratch"*, DeepMind blog, October 2017, http://bit.ly/2zQaEQW

SPONSORED BY

www.teradata.com

TERADATA.