kinetica

BI Meets AI

Augmenting Analytics with Artificial Intelligence to Beat the Extreme Data Economy

Inside

WHITE PAPER

Executive summary

Many business leaders struggle to realize the financial impact of investments in data lakes and other big data initiatives. Collecting and storing huge volumes of business data is the easy part. Finding new sources of competitive advantage in that data is a decidedly harder task.

It's getting harder. We now live in a world where the economy runs not on services, but on data itself. It's not about which or how much or how quick the business can collect the data anymore. In the Extreme Data Economy, data comes from unpredictable sources and analysis becomes much more complex. Data can be big or small, static or streaming, structured or unstructured, human or machine. On top of that, data's useful life may only be hours or minutes, or perhaps weeks or months; the method of extracting value from data may be different depending on its lifespan. The legacy databases built for serial computing and data lakes just can't keep up.

Artificial intelligence (AI) is emerging as a much-needed catalyst for extracting value in data because it accelerates the rate and increases the depth of analysis. Leader organizations are now using GPU parallel processing technology to infuse AI into their Business intelligence (BI) applications, and this strategy is quickly defining the next generation of business analytics. This whitepaper defines the key requirements for an insight engine that can support artificial intelligence, and the ways businesses can optimize analytics with AI in order to gain a competitive advantage.

Glossary of Key Terms

CPU: Central Processing Unit. A common computational chip that carries out instructions, it indexes and uses batch-processing to analyze data linearly.

Batch Processing: A non-continuous mode of running data queries, which requires more time than parallel processing but is often required when computational power is limited.

Indexing: A time-consuming process in which the contents of a database are abstracted into a more digestible structure; required for all non-GPU databases.

GPU: Graphics Processing Unit. The most powerful computational chip available, it is able to ingest and analyze data in parallel, and is therefore the fundamental compute device for machine learning.

Machine Learning (ML): The subset of artificial intelligence most commonly used in data analytics.

Data Locality: A key attribute of next-generation analytics, when data required for computation is stored as close as possible to the point of computation.

In-Memory Database: IMDB stores data in a computer's main memory (RAM) instead of a separate storage device, enabling exponentially faster computation.

UDF: User-Defined Function, an attribute of next-gen analytical databases that enables business users to perform AI-like functions inside the database.

Data Bottlenecks: A point at which the business need for data outpaces the ability of IT to deliver.

Insight Engine: A next-generation, unified artificial intelligence and business intelligence analytical database that utilizes both in-memory storage and GPU processing. It depends on ML algorithms to reveal potentially valuable patterns in data.

1 / The problem with most big data projects: batch processing and compute bottlenecks

For all the excitement around big data, most businesses are still struggling to realize business value from their data lakes and Hadoop deployments. A 2017 Forrester survey of 2,094 enterprise decision makers revealed that advanced analytics and more performant data architecture occupies 4 of the top 5 spending priorities. The age of big data — largely defined by complex platforms like Hadoop — has entered Gartner's fabled "trough of disillusionment" where inflated expectations meets the reality of imperfect technology. A Gartner study in 2017 found that only 14% of Hadoop deployments are actually in production.

Storing data, it turns out, is a lot easier than finding insights. Two critical factors limit the value of big data in most businesses: the depth of analysis and the speed of analysis. Until Artificial intelligence arrived, going deep into big data required the rare talent of a data scientist who knew what questions to ask and how to harness the data. Even with AI, data scientists continue to play an important role. Getting fast results from even the smartest queries is complicated by the fact that most big data architecture still depends on batch processing using CPUs instead of parallel processing powered by GPUs.

If you've ever spent minutes (or hours) waiting for your chart to display in a Business Intelligence (BI) application like Tableau or Power BI, you understand how slow analytics that run on traditional databases can be. In the popular Hadoop ecosystem, Spark is often described as a "real time" analytics engine but in reality Spark is often relegated to batch processing because the database managing the data simply can't keep up. This means analyst's queries are answered by a series of relays between the application asking the question and a separate data store that contains an index of the information needed for the answer. The faster the speed of the data coming in, the longer it takes for index in the database to update and the longer it takes for query results to be sent back to the application. Even when data and processing tasks are distributed across multiple machines to optimize speed, the process is relatively slow compared to new data architectures that hold data in memory instead of disk and use GPUs instead of CPUs for processing.

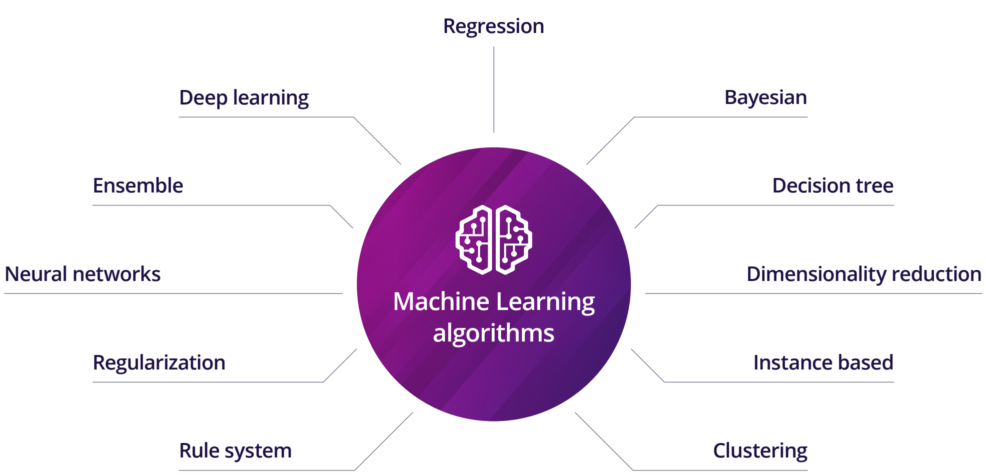
> "
> This is about accelerating time to insight. GPU and in-memory data make the results of big data, AI, and Machine Learning far more accessible inside business processes, because you can get answers in time."
>
> **Dan Woods**, Forbes writer

This new approach to hardware acceleration is designed for the Extreme Data Economy. It not only exponentially speeds up draw time for business dashboards, but also opens a new set of opportunities for deep analytical insight powered by artificial intelligence. Combining AI with BI is now a real option, and the results are impressive. "This is about accelerating time to insight," said Forbes writer Dan Woods during a recent webinar hosted by Kinetica. "GPU and in-memory data make the results of big data, AI, and machine learning far more accessible inside business processes because you can get answers in time."

2 / What is artificial intelligence and machine learning?

Artificial intelligence (AI) is a broad category that encompasses the many ways machines can be used to execute tasks in ways that people consider smart. AI has been around for decades in various forms. AI is often equated with the kind of deep analytical work seeking insights in big data. Machine learning (ML) is a more precise way to describe how AI is applied to big data. Machine learning is a practice area in AI that essentially allows computers to find patterns in data without being explicitly programmed to solve a single problem. The work ML does when you point it at a large data set is driven by one of a dozen ML algorithms. Some of the more popular ML algorithms include deep learning, regression, clustering, neural networks, and Bayesian analysis. Unlike the way a human approaches a dataset, ML doesn't go into the data knowing the problem it is trying to solve. ML algorithms simply find and predict patterns, and that's when teams from business, data science, and analytics become part of the equation. Humans ultimately need to take the baton from a ML algorithm, interpret what the patterns mean, and determine if there's business value in understanding or changing the pattern.

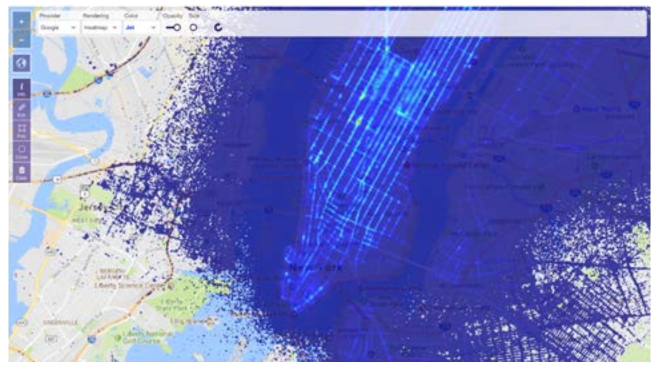
Financial services largely pioneered the commercialization of machine learning, but several industries are now capitalizing on this special kind of AI because computing power and database software have finally evolved to the point where businesses can afford to develop their own ML capacity. Some of the more common use cases of ML in business today include predictive modeling and natural language processing. Think about Siri, Amazon Echo, or Cortana. ML allows these digital assistants to evolve and adapt to the way you speak and the things you need. And the use cases go way beyond consumer technology. Cognitive applications in healthcare, retail, manufacturing, and transportation are saving lives, improving business efficiency, and delivering new kinds of customer experiences that feel more personal and relevant.



3 / How BI + AI accelerates data insight

Adding artificial Intelligence (in the form of machine learning) into business intelligence is the most impactful way to accelerate data insight. Establishing a unified AI+BI database—an insight engine—means a business organization can move from an analytics posture that looks back to one that looks forward. This is the difference between asking "what happened in the past?" versus asking "what will happen in the future?"

There are multiple use cases emerging from businesses that combine AI and BI in one of these next-gen analytical insight engines that use both in-memory storage and GPU processing. Retailers are transforming supply chain management, for example, as they can now ingest and analyze streaming data from suppliers and shippers against real-time inventory data from retail operations. This new capacity for mass streaming ingest combined with concurrent high-throughput analytics is giving leading retailers the ability to make real-time adjustments in the supply chain based on data in motion, including point of sale data, weather data, social media data, and more. This creates a competitive advantage because critical supplies are re-routed where they're likely to be needed much sooner than ever before. Retailers are also combining AI and BI to find new affinities between products and customers, resulting in higher performance from recommendation engines.



Mail delivery services can use next-gen analytical insight engines with GPU processing and in-memory storage to optimize the efficiency of delivery routes in a way human minds could never achieve.

Logistics is also capitalizing on the power of next-gen analytics. As data related to locations continue to grow with the internet of things (IoT), logistics businesses are finding new ways to visualize and optimize the interactions between people, products, and things. Geospatial analytics is one of the most demanding workloads because of the high volume and variety of IoT-related data streams and the immense computational power required to visualize data on a map. Mail delivery services can use next-gen analytical insight engines with GPU processing and in-memory storage to optimize the efficiency of delivery routes in a way human minds could never achieve. While use cases vary for unified BI + AI insight engines, they usually share two attributes. They leverage the incredible computational power of the GPU instead of the CPU, and they depend on a machine learning algorithm instead of humans to reveal potentially valuable patterns in data.

4 / Key requirements for a next-gen analytical insight engine

In addition to the fundamental technical requirements of GPU and in-memory storage, next-gen analytics requires executive leaders to rethink organizational workflows to **combine BI and data science teams**. With GPUs and in-memory technologies finally at the ready, businesses beginning their journeys into machine learning will likely face challenges overcoming legacy operational models that separate these two groups. It's quite common for IT organizations to deploy separate databases for BI and ML projects. This is a natural reflection of the way data science and ML teams are typically segregated from BI departments. ML and data science teams tend to work on different hardware, work independently from the data warehouse (creating new data silos), and most critically, tend to work in isolation from the business subject matter experts in BI. For BI and AI to work together, organizations must overcome the process silos between separate hardware, separate platforms, and separate teams.

When it comes to choosing a next-gen analytical insight engine to support your BI + AI program, **SQL support** is a critical requirement for extending existing investments in software and talent. It allows for the seamless connection of existing analytic apps, and enables IT to operationalize machine learning using the team's existing skill sets. SQL is the cornerstone of traditional business databases, driving integration across the IT stack and enabling business users, analysts, developers, data scientists, DBAs, and system architects to access and analyze data. A next-gen analytical insight engine like Kinetica gives diverse teams common access with built-in SQL support.

An enterprise-grade insight engine should also provide an **API** (Application Programming Interface) so more advanced developers and data scientists can have richer programmatic access to the database from within their custom apps. Similarly, it's important that your analytical insight engine support the full ecosystem of tools used in today's cobblestone world of analytic apps; this requires a full set of connectors for Apache Hadoop, NiFi, Spark, Storm, Kafka, and commercial tools like FME. Business analysts will require ODBC/JDBC connectors for BI tools like Tableau, Microstrategy, and Power BI.

Scale-out architecture and **high availabilit**y are critical as well. Scale-out architecture means your insight engine can distribute the information and workloads across multiple machines after you max out the available GPU slots on your first machine. High availability is an important feature because it ensures that data remains accessible—and thereby, business continuity—if one machine in your scale-out architecture fails.

Most businesses are still using serial computing running on CPUs to store, manage, and analyze data. Problem is, these technologies are just too slow to extract value from data in real-time. An insight engine powered by GPUs can take in and analyze data in parallel, in real-time, unlike technologies built for CPUs. It can perform **real-time analytics on streaming data**.

An insight engine should dramatically accelerate analysis on billions of rows of data, with in-memory GPU architecture using **accelerated parallel processing**. Expect it to deliver results in milliseconds. It should provide near-linear scalability without the need to index.

Ideally, an insight engine helps you discover **visual insights** as well. It can even take geospatial and streaming data and turn it into visualizations that reveal interesting patterns and business opportunities, capitalizing on the GPU's particular aptitudes, including rendering the visuals themselves.

Finally, evaluate any next-gen analytical insight engine to ensure it provides **built-in machine learning** capabilities. For example, will it run Google's popular Tensorflow and other AI frameworks via User Defined Functions that analyze the complete data corpus? Allowing the insight engine to manage the frameworks means trained ML modules can be used by business analysts on other datasets, which moves the business closer to a fully operationalized ML capability—allowing even non-technical users to access ML within popular apps and reporting tools like Excel, Tableau, TIBCO Spotfire, and Microstrategy.

5 / Starting the BI + AI conversation with executive leadership

"There's a real mindset in the enterprise right now that ML can change the way they do business," said Kinetica CTO Nima Negahban. But with most IT departments occupied managing infrastructure, it falls to profit-driven business units to lead the discussion around implementing an insight engine for BI + AI. One of the easiest way to start the conversation, according to Kinetica Chief Strategy Officer, Amit Vij, is to focus on the speed of business data. "When we're talking to CEOs they're saying, 'I'm looking at reports that are a week old. If I had this data a week ago, we could have done something that made a difference.' A lot of people don't understand that it's not unrealistic anymore to expect critical business data in near-real-time."

According to Vij, business leaders should ask three related questions to spark the conversation about a next-gen analytical insight engine capable of supporting BI + AI:

1 / Where are the data bottlenecks? Find out where the choke points are in your enterprise data architecture. Explore the kinds of decisions that are delayed because business data is hampered by processing time or availability. Both of these barriers are resolved with GPU processing and in-memory storage, the essential hardware of next-gen analytics.

2 / How much potential value could we capture if we removed our data bottlenecks? What business processes would improve if data was available in near-real-time? What are the potential financial impacts of making key decisions faster?

3 / If we piloted a next-gen analytical insight engine to solve for BI speed, where could we focus a POC in AI? Standing up the infrastructure to accelerate BI to near-real-time delivery will be a win in and of itself. But that same infrastructure opens the door to your first machine learning proof of concept. Which data sets would be most interesting to feed in? Which ML algorithms are most fruitful with data like yours?

Priority Requirements

Business executives should look for the following elements when selecting a unified AI + BI analytical insight engine.

- Combine BI + Data Science Teams
- SQL Support
- REST API
- Data Persistence
- Scale-Out Architecture
- High Availability
- Real-time analytics on streaming data
- Accelerated parallel processing
- Visual insights
- In-database machine learning

One of the easiest ways to start the conversation is to focus on the speed of business data.

kinetica

The journey to a fully operationalized AI capability that can withstand the requirements of the Extreme Data Economy begins with a single step toward the right infrastructure. Browse our solutions to explore the transformative impact a next-gen analytical insight engine could have on your business. Get in touch to set up a solutioning session with a Kinetica expert who can help bring AI to your BI. Or get a demo to see Kinetica in action.

Explore Solutions

Get Demo

About Kinetica

When extreme data requires companies to act with unprecedented agility, Kinetica powers business in motion. Kinetica is the instant insight engine for the Extreme Data Economy. Across healthcare, energy, telecommunications, retail, and financial services, enterprises utilizing new technologies like connected devices, wearables, mobility, robotics, and more can leverage Kinetica for machine learning, deep learning, and advanced location-based analytics that are powering new services. Kinetica's accelerated parallel computing brings thousands of GPU cores to address the unpredictability and complexity that result from extreme data. Kinetica has a rich partner ecosystem, including NVIDIA, Dell, HP, and IBM, and is privately held, backed by leading global venture capital firms Canvas Ventures, Citi Ventures, GreatPoint Ventures, and Meritech Capital Partners. For more information and trial downloads, visit **kinetica.com** or follow us on **LinkedIn** and **Twitter**.

kinetica

Headquarters

101 California St, Suite 4560
San Francisco, CA 94111

(888) 504-7832
+1 415 604-3444

To learn more

visit us at
kinetica.com

or
contact us