Next-Generation Predictive Applications

December 4, 2017

Rick Sherlund
Partner, Co-Head of Technology
(415) 671-4557
rsherlund@pwpartners.com
Executive Summary

• An upcoming platform shift in computing will enable a new generation of predictive applications that leverage the enormous growth in big data and widespread availability of machine learning (ML).

• Data is the critical new strategic asset, driving ML algorithms that are essential to competitive differentiation across all industries.

• This shift will be expansive for the software industry, but potentially challenging to existing SaaS vendors, who until now have architected on traditional, scale-up architectures and relational database platforms.

• The aggregation of big data sets and democratization of artificial intelligence (AI) and ML in the cloud will play to the strengths of the leading public cloud platform providers, who will gain increasing scale advantage in the tech sector. This leading cloud market position could be used to migrate up the technology stack, allowing Amazon, Microsoft and Google (among others) to compete more directly in the infrastructure/middleware software space and potentially to consolidate up into the enterprise applications space.

• A significant number of new business processes will be automated, driving a renaissance in business process reengineering and low-code business process management (BPM), as well as the need for cloud-native middleware including tools for implementing a microservices architecture, containers and orchestration.

• The value of the next generation of applications will not be so much about how they are programmed to automate workflows, but rather the new processes that will need to be automated and the value of the logic/algorithms that will be synthesized from ML to enable optimization and real time decision making.
The Next Generation of Predictive Applications: The enormous growth in data and the ability to economically aggregate big data sets in the cloud and to apply AI and ML to extract predictive insight from this data will drive innovation and enable an emerging new platform of enterprise computing and a new generation of predictive applications.

These technologies will leverage the widely available public cloud computing infrastructure (Amazon AWS, Microsoft Azure, Google Cloud Platform as well as a dozen other leaders, including IBM, Oracle and Alibaba). Public cloud computing platforms have evolved over the past 10 years, but only recently have these platforms been leveraged to address the ability to economically aggregate big data sets in the cloud, in close proximity to the vast computational resources of the scale-out public cloud platforms, and apply recent advances and widespread availability of ML to derive predictive insights from these enormous data sets.

Even the leading SaaS vendors, who disrupted the previous generation of on-premise systems, have engineered on this legacy platform of computing. The SaaS apps are mostly architected using multitenancy and are deployed as-a-service in the cloud, but are still fundamentally transaction process and workflow automation systems architected on the traditional relational database management systems (RDBMS) and scale-up platforms.

The architecture underlying these systems was not designed for and is not capable of supporting the enormous unstructured big data sets that are emerging and the near limitless scale-out compute resources of the public cloud infrastructure. The new platform is required to support ML and derive the predictive algorithms that define the next generation of applications. This platform and new generation of predictive applications will distinguish a digital-first competitor from a legacy enterprise.

Digital-native companies do not have the baggage of legacy systems. We envision the need to replace or augment the current generation of transaction systems that run enterprises with a new platform of computing and a new generation of predictive applications that leverage big data and machine learning on top of the public cloud infrastructure.

The next generation of predictive systems will be both market-expanding and potentially disruptive for the software sector. There are also broader, macro and geopolitical implications for workers and societies from the next wave of automation and application of AI to reduce labor and operating costs, optimize production and distribution and, at a macro level, enhance economic productivity for the countries that lead in the transition to new digital technologies.

The platform shift will also be a further, broader catalyst for the digital transformation of businesses across all industries that are starting to see the strategic value in customer data and are fearful of digital-native competitors and of being “Amazon’d,” “Uber’d,” Airbnb’d” or “Tesla’d.”

Digital transformation is well underway across all industries. The progress is constrained, however, by the limitations of the current software platform running traditional enterprises. Transactional systems (so called systems of record) have been built on an older, scale-up architecture running RDBMS technology. These systems are efficient in automating workflows and scaling up in performance to record high volumes of transaction details (transaction processing), but are not suitable for addressing unstructured big data sets and ML.

We focus in this report on the implications for the software industry. This will lay the foundation for further discussions to follow on the broader implications of digital transformation across industries, in collaboration with our colleagues here at Perella Weinberg Partners who have the respective industry specific expertise.

The digital transformation of production and distribution in the manufacturing and energy sectors has centered on the internet of things (IoT), which is a good example of the predictive capabilities of the new platform, so we address this at a high level. For a deeper dive into IoT and Industries 4.0 we will need to leave deeper insights into this market for a follow-up discussion.
Data is the new strategic corporate asset driving the platform shift. With more digital devices in the hands of consumers (smartphones) and sensor enabled devices across industries we are witnessing a proliferation of data sources and data types (Figure 1).

Data is coming from all directions, from web click streams, geolocation data from smartphones, video cameras, billing records, call details, electronic medical records, genome data, census and other public records, weather data, sensor devices from the industrial and energy sectors and home automation systems, log streams, social media, ecommerce data etc. This data is referred to as unstructured data. Transactional data that is stored in rows and columns of a RDBMS is structured data and comprises the systems of record of an enterprise. The growth in unstructured or big data is more of interest in extracting predicative insight that can drive real time decision making for the enterprise and transform business processes.

**FIGURE 1: DATA VOLUMES GROWING FASTER THAN MOORE’S LAW**

With the rapid growth in the ability to collect this data (“big data”) and store it economically in the public cloud platforms (or similar large scale private cloud platforms), we can now begin to apply the rapid advances in AI and ML to act upon this data, in real time, and extract insights that can inform a new generation of applications. This requires enormous compute resources and a new architecture which consumer internet companies (Google and Facebook) pioneered and is widely available in the public cloud platforms to be leveraged in support of the aggregation of big data and the application of AI and ML to uncover the algorithms that express the relationships among the data for predictive insights.

The rapid decline in prices of storing data is a contributing factor in the ability to aggregate these big data sets. Ten years ago it cost about $0.80 to store a GB of data. This is now headed down to about $0.005 (half of a cent) (Figure 2). (Note that Figure 2 shows the dramatic decline in storage costs on a log scale.)

---

**Call Detail Records**  **Click stream**  **CSV**  **Data**  **Documents**  **Emails**

**Billing Data**  **Meta Data**  **JSON**  **Mobile Data**  **Network Data**  **PDF**

**Product Catalog**  **Medical Records**  **Text Files**  **Text Messages**  **Video**  **XML**

**Sensor Data**  **Server Logs**  **Set Top Box**  **Audio**  **Social Media**  **Merchant Listings**

---

Source: McKinsey Global Institute
**FIGURE 2: SUBSTANTIAL DECLINE OF HARD DRIVE COST PER GIGABYTE (USD)**

For the software industry, this represents a game changing new platform of computing. We will see new user interfaces that leverage AI, natural language technology and virtual reality/augmented reality. This is already being incorporated into chat bots for customer interaction and in personal digital assistants, including the Amazon Echo, Apple Siri and Microsoft Cortana. We will also see the enablement of a new generation of intelligent and predictive applications and real time business processes that act on the insights derived from the ML models.

Existing SaaS industry leaders have almost all architected on the traditional, scale-up architecture and transactional relational databases. Disruptive new vendors likely will capitalize on the limitless scalability and big data storage capacity of the scale-out public cloud platforms first used by consumer web-scale vendors such as Google and Facebook, who have proven the robustness of the architecture. The public cloud platform has been in existence for 10 years now, but it is the recent addition of big data sets and the democratization of ML on this platform that leverages this architecture to enable next generation systems.

Transactional systems have historically been used to automate workflows, business processes and record keeping (systems of record) and, more recently, with front office systems for customer engagement (systems of engagement). We believe the next generation will be about systems of prediction. Transactional systems will need to be adapted to integrate with real time systems of prediction or themselves be rewritten to become predictive systems.

We would expect the next generation of transactional systems to benefit from a rejuvenation and replacement cycle with predictive capabilities and real time processes. The concern we have is that existing platforms were not architected for the enormous volume, velocity and variety of data being collected and the compute power required to act upon this data to train the ML engines to derive the predictive algorithms required to be competitive in the new digital enterprise. There is a need for a new computing platform that can address these requirements. The leading SaaS vendors, who were
AI and Machine Learning: Artificial intelligence (AI) is a comprehensive term. The technology is years away from being as broad based as human intelligence, but rather there are narrow applications of the technology that are being used such as in natural language processing for personal digital assistants (Amazon Echo with Alexa, Apple Siri, Microsoft Cortana) and speech to text translations (Nuance was a pioneer for medical record transcriptions) and used in enterprise software for automated customer service (chat bots from Salesforce for example) and for facial recognition (Apple iPhone X) and other image recognition (autonomous vehicles).

IBM’s highly publicized Watson can be trained with text input and demonstrate knowledge of specific topics it has been trained in (supervised learning). For predictive insights, however, we are seeing the widespread adoption of ML. ML is a subset of AI and is math based and can train itself (unsupervised learning) to recognize patterns, relationships and correlations. For example, ML is beginning to be used for predictive maintenance in IoT applications. Machines are being sensor enabled to stream data on pressure, temperature, vibrations, etc., which can be used for predictive capabilities (predictive maintenance or asset performance management) once the relationships have been understood and
expressed in the models or algorithms. Predictive capabilities have also been used in recommendation engines from Amazon and Netflix and more widely in marketing automation systems from Adobe and Salesforce.com.

Training a ML engine requires large amounts of complex data and enormous computational power, which has not existed economically until recently. To derive optimal predictive capabilities these algorithms can be refined over time as the models “learn” (are trained) from greater usage when they have been deployed in enterprise systems.

Deep learning is a branch of ML which has more layers to better refine and optimize resulting predictions or classifications. For example, the difference between a dog and a cat (or the difference between a dog and a pedestrian for autonomous vehicles) may require more layers of complexity, analysis and compute capacity to sub-classify and discern differences. We will mostly refer to ML, but the requirements may be for even more complex deep learning systems as well.

ML engines are widely available as open source tools, many from universities or contributed by vendors such as Google’s popular TensorFlow library of ML algorithms, Amazon Machine Learning or Azure Machine Learning. Others are offered by specialized vendors, many of which focus on a specific vertical markets and combine their ML technology with professional services with the domain expertise of the market they are addressing. Many of these small companies are being acquired (“acquihires”) as various industries attempt to build their knowledge and expertise in this area. The public cloud vendors also host their respective ML engines as-a-service and have broader AI technology such as image recognition available through API’s on their platforms, leading to widespread availability and democratization of AI and ML.

For performance purposes it is important to have the big data sets in close proximity to the computational power and to the ML engines that act upon this data. The implication being that scale advantage accrues to the public cloud vendors who provide low cost storage and nearly limitless horizontal, scale-out compute resources. The public cloud vendors have scale advantage and have been active in acquiring AI and ML technology which they now provide as-a-service (Figure 4). The widespread availability, or democratization, of ML is a contributing factor to the emergence of the new generation of systems and predictive applications.

The ML engines find non-obvious relationships that are typically beyond the ability of humans to comprehend given the number of dimensions and volume and variety of the data, and produce the algorithms that express these relationships. This is referred to as training the ML engines.

Once these algorithms are built they can be deployed and integrated into the new computing infrastructure stack. Their predictive capabilities can then be used to inform the applications layer, enabling a new generation of predictive, intelligent and real-time applications. It is this new class of intelligent apps that will alter business processes across the enterprise and challenge existing enterprise applications. There will also likely be new user interfaces for all classes of applications that leverage natural language processing and in some cases virtual reality and augmented reality.
### FIGURE 4: AI-as-a-SERVICE OFFERING OF CLOUD PROVIDERS

<table>
<thead>
<tr>
<th>ML Service</th>
<th>Amazon Machine Learning</th>
<th>Azure Machine Learning</th>
<th>Google Cloud Machine Learning</th>
<th>Watson Machine Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Recognition</td>
<td>Rekognition</td>
<td>Computer Vision API</td>
<td>Vision API</td>
<td>Visual Recognition API</td>
</tr>
<tr>
<td></td>
<td>Kinesis</td>
<td>Content Moderator</td>
<td>Cloud Video Intelligence API</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Emotion API</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Face API</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Video API</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Language Processing &amp; Speech</td>
<td>Lex</td>
<td>Bing Entity Search API</td>
<td>Natural Language API</td>
<td>Alchemy Language API</td>
</tr>
<tr>
<td></td>
<td>Polly</td>
<td>Bing Spell Check API</td>
<td>Translation API</td>
<td>Conversation API</td>
</tr>
<tr>
<td></td>
<td>Transcribe</td>
<td>Language Understanding Intelligent Services</td>
<td>Speech API</td>
<td>Document Conversion API</td>
</tr>
<tr>
<td></td>
<td>Translate</td>
<td>Linguistic Analysis API</td>
<td></td>
<td>Language Translator API</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Text Analytics API</td>
<td></td>
<td>Natural Language Classifier API</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Translator API</td>
<td></td>
<td>Personality Insights API</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Web Language Model API</td>
<td></td>
<td>Retrieve and Rank API</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bing Speech API</td>
<td></td>
<td>Tone Analyzer</td>
</tr>
<tr>
<td>Prediction</td>
<td>SageMaker</td>
<td>Custom Recognition Intelligence Service</td>
<td>Speech to Text/Text to Speech</td>
<td>Speech to Text/Text to Speech</td>
</tr>
<tr>
<td></td>
<td>DeepLens</td>
<td>Speaker Recognition API</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Sumerian</td>
<td>Academic Knowledge API</td>
<td>Google Cloud Jobs API</td>
<td>Watson Discovery</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Entity Linking</td>
<td></td>
<td>Watson Knowledge Studio</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intelligence Service</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Knowledge Exploration Service</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Perella Weinberg Partners

**Public Cloud Platforms as Enabling Infrastructure:** Hyper-scale, public cloud platforms now make it economical to store enormous amounts of complex data or so called “big data” data sets. Big data is not just a lot of data, but very complex data sets that need to be correlated with many other very large data sets, such as sensor telemetry, weather data, census or other public records, including traffic for smart cities, log files, e-commerce, social media and web click streams, etc. Often these relationships are multi-dimensional and involve dozens of different data sets to uncover non-intuitive (well beyond human comprehension) relationships among the data.

The enormous storage, networking and compute resources required for this new generation favors the superior economics of public cloud platforms such as Amazon AWS, Microsoft Azure and Google Cloud Platform (GCP) (followed by a long tail of about 10 others, including IBM, Oracle and Alibaba) where workloads can be scaled-out across hundreds of conventional processors (CPUs) in parallel for near limitless scalability.

Training ML engines is highly computational intensive and can strain even the scalability of public cloud platforms, so specialized processors such as GPU’s from Nvidia or TPU’s from Google are now being hosted as-a-service in the public clouds to address the growing market requirements.

The new platform of enterprise computing is likely to be cloud native, with the advantage of being able to scale the compute and storage out (scale-out architecture) across tens or even hundreds of processors operating in parallel in the cloud. Current SaaS vendors mostly use the traditional scale-up architecture, which requires larger and larger computers or engineered systems and do not easily address big data requirements of the magnitude we believe enterprise will need to address. The market is likely to move to a cloud-native architecture, following the lead of the large web-scale companies,
which have proven its scalability in demanding consumer internet markets.

It is the I/O (pathway for getting data in and out of the CPU) of scale-out systems that enables much of the scalability for big data sets. The centralized I/O of scale-up systems is a bottleneck for performance and scalability; it is removed as a chokepoint in scale-out architecture, with each processor, in parallel, having its own I/O to deliver linear scalability. The extreme demands for scalability for the early consumer internet companies (Google and Facebook) resulted in the need to invent the scale-out architecture, which has now moved from the consumer web-scale companies to the cloud-native architecture at public cloud vendors AWS, Azure and GCP for use in enterprise computing.

**Public Cloud Vendors Could Disintermediate Middleware Vendors and Consolidate SaaS.** With the ability to now aggregate and store big, complex data sets economically in the cloud and with the computational power and ML resources also hosted as-a-service, we would expect the public cloud platforms to increase their scale advantage as a platform for the next generation of enterprise systems. This would enable the public cloud vendors to build upon their platforms and scale advantage to move further up the infrastructure and middleware stack to potentially disintermediate infrastructure software vendors (offering their own middleware or open source infrastructure components such as the Apache ecosystem) and potentially to consolidate further up the technology stack into the applications space.

This would be similar to prior technology cycles where the systems software vendors (IBM, Oracle and Microsoft) leveraged their foundation or underlying platforms for the technology benefits of product integration and the economic benefit of leveraging customer relationships and an existing direct sales force to consolidate up the stack into middleware and applications.

The applications (referred to as Software-as-a-Service or SaaS) layer would be highly strategic for the public cloud vendors to consolidate. This could serve as on-ramps to fuel the provisioning of networking, compute and storage (referred to as Infrastructure-as-a-service or IaaS) and higher lever middleware for building applications (referred to Platform-as-a-Service or PaaS). By consolidating into the applications layer, this might also pull along the compute and storage of the broader ecosystem of customers and partners of the enterprise apps vendors.

Also, having ownership of apps would imply owning the ML algorithms derived from the big data that is relevant in these markets. The quality of the predictive algorithms could be an important competitive differentiator and defensive moat for the apps participants. Apps and the related ML algorithms would be “sticky” (tending to retain customers) and drive public cloud IaaS and PaaS and perhaps the ecosystem of partners and a new generation B2B or B2C marketplaces.

Horizontal (cross-industry) applications, such as sales automation, human capital management (HCM), financials and broader ERP (enterprise resource planning) make the most sense to be offered by the public cloud vendors, perhaps initially for smaller and medium-size companies, while vertical market applications are a very fragmented set of applications that require much greater industry-specific expertise and specialized sales resources. Some vertical markets may be more strategic to a cloud vendor than others.

The decision to move into the applications layer requires an enterprise sales force that can call on departmental users of the applications and ongoing customer support, versus the infrastructure layer, where one would typically sell to and support IT in a more cost-effective, leveraged fashion.

The applications software market is highly fragmented, with horizontal apps in HCM, finance and ERP to industry specific apps across dozens of different industry sectors. Even with consolidation into the apps space with some segments dominated by integrated public cloud stack vendors, this would likely remain a highly fragmented market.

**Risk to Existing SaaS Vendors:** This new generation of predictive applications presents a risk to existing SaaS vendors, who have architected on traditional scale-up computers and transactional relational databases, which are not as well suited for handling the enormous I/O and processing requirements of complex big data sets and ML.

The ability to extract predictive insights from complex big data sets, which will influence all applications...
in the future, is best addressed by leveraging the newer scale-out architecture and big data databases supported as-a-service in public or private cloud platforms.

We have considered whether this upcoming platform shift would be a technology discontinuity in the market which could severely hinder and disrupt existing SaaS vendors. In historical platform shifts, leading vendors are often replaced by emerging new vendors who can fully embrace the new platform. Sometimes these shifts are more evolutionary and can be managed with a bridge to the new platforms and therefore market expanding to the entrenched vendors.

For the class of applications that are more workflow automation driven and might offer less benefit of predictive capabilities around big data, it may be possible to bridge over in a more evolutionary progression using a hybrid approach that integrates well with the new predictive generation of apps without rewriting and losing the benefits of transactional integrity of the RDBMS architecture. Other more big data influenced apps would likely requiring a rewrite.

This, of course, is very situation dependent, based on the size of the data sets and the value of predictive capabilities for each respective application. The traditional scale-up and RDBMS architecture is unlikely to adequately address the new class of predictive and intelligent applications that must support the explosive growth in data (volume, velocity and variety) to deliver real-time systems such as digital marketing, e-commerce and IOT applications. Predictions are only as good as the volume of data that enable them, so it is the increased ability to aggregate large data sets economically in the cloud and to drive digital transformation across industries that will fuel this expanding new class of applications.

Apps vendors must be thoughtful about how to adapt to the need of the next generation of intelligent and predictive systems and not just bolt-on some analytics to existing systems or apply older rules based systems in order to “check the box” for marketing purposes and assert that the apps is a ML based system.

We believe most of the existing SaaS vendors will need to adopt a hybrid approach, likely preserving their existing investment in transactional systems, but bridging over to public cloud platforms where the need for big data and ML is required to deliver greater predictive insights. This is a complex issue that is just now being addressed by vendors who likely see the need for predictive capabilities but also see the limitations of their existing architecture.

One might consider training the ML models in the public cloud (this could represent many TB’s or PB’s of training data) but deploy the algorithms on a conventional scale-up architecture if the streaming data volumes were not too demanding. The high end of the engineered systems (scale-up systems) such as SAP's Hana can address up to 1 TB of data. This is small compared to the standards of big data. The need to train models typically addresses much larger data sets and may require specialized GPU's and scale-out capabilities, but once the models are created they may, in some cases, be deployed on a traditional scale-up architecture. Applications, however are built on top of database management systems and are difficult to modify without rewriting the application. Some architectural re-engineering must be done to bridge the transactional systems with the next generation of predictive systems and there is much work going on in the industry.

The transactional applications are mostly architected on RDBMS from leading vendors including Oracle, Microsoft and IBM. RDBMS are highly structured with each record as a row and each column an attribute of that record and are queried using a structured query language (SQL). The structure or schema is fixed and managed by database administrators and this structure helps optimize read and write performance across many tables of records. RDBMS have a high level of performance and transaction integrity, but are not well suited for unstructured data.

Big data sets are unstructured data and are better stored in big data databases. These might include No-SQL databases such as the document database MongoDB or MarkLogic or the time series or key value stores such as the DataStax open source Cassandra database or Couchbase or the Apache Hadoop distributions like Cloudera and Hortonworks. Amazon offers its own Dynamo DB and Microsoft offers Cosmo DB on Azure as NoSQL databases suitable for unstructured data types. These may not
offer the transactional integrity of a traditional relational database so there is an emerging need to have
the capabilities of both in a single platform (MapR has a “converged” database) and there are broader
efforts underway to address this.

A hybrid approach may be possible as the industry works to bridge the two platforms. We are seeing
some rewriting of systems for the new platform that requires a high level of predictive insights such as
the Salesforce marketing automation system.

We question the adequacy of existing platforms to address the next generation of predictive systems
but must allow for the ability to engineer hybrid solutions that may be a bridge, with the most
demanding apps having to be rewritten and most likely hosted in the public cloud. For a startup
vendor, it is clear that one would leverage the public cloud platforms and build on top, leveraging big
data storage and ML capabilities to arrive at next generation predictive apps hosted in the cloud. We
have been watchful for this new generation of companies and we are just now beginning to see some.
These companies also offer more modern middleware supporting a microservices architecture and the
use of containers and Kubernetes for automating deployment, scaling and management of containers.

As a final point on the architecture, there may be confusion over vendors stating they are supporting
AWS (or other cloud platforms) by moving traditional apps over to this platform. To be clear, this is
a re-hosting onto someone else’s computers, like the traditional ASP (Applications Service Provider)
or a managed service provider models. To gain the advantage of a scale-out platform and big data,
the application needs to be rewritten not simply “fork-lifted” over to the public cloud. The traditional
relational database architecture does not address unstructured big data in a technical or an economical
way and the application does not take advantage of the enormous compute capacity of scale-out
computing. So hosting on AWS, Azure or GCP does not make a traditional app a next generation app.

Digital Transformation, Real-Time Systems and Strategic Value of Data:

Front office or customer-facing applications (sales, marketing and support) were early drivers of the
digital transformation, and hence are an important market for the new systems of prediction. Digital
transformation often starts with the front office, or customer interaction.

The back office ERP applications are more likely to be driven by IoT and the need for systems to
accommodate the streaming of sensor telemetry for real-time insights that can impact production,
logistics, billing, etc. The industrial sector early on harnessed large-scale sensor data streaming in real
time and at high volume and velocity for predictive maintenance and asset performance management.
The IoT stacks are excellent examples of systems of prediction, so we take a high level look at this in
this report.

Data is the strategic competitive advantage over traditional systems, which historically have been more
about the software to automate business processes and workflows. In the future, companies will not
buy software (SaaS) just for automating processes and workflows, but rather to gain a competitive
edge by leveraging the value of data via higher-quality algorithms/predictions. The value of the
new software platform is that it enables predictive capabilities from complex data sets through the
application of AI and ML. Software that does not embrace predictive capabilities and drive real-time
insights will be an impedance mismatch with the real time digital enterprise and will need to adapt.

The modern cloud native architecture leverages open-source technologies, including Spark streaming
and Kafka messaging, which recognize the time value of data and transform computing into real-
time systems and transforms companies into real-time enterprises. Predictive applications and real-
time systems are complementary and together drive real-time and automated decision making. They
also impact business processes (real-time processes) and drive the need for integration across the
enterprise to orchestrate front-office and back-office functions.

The move to predictive applications and real-time applications may be an incentive for consolidation in
the software sector. This catalyst will derive from the need to create and leverage a common customer
data record across the enterprise, transcending existing best-of-breed siloes as companies seek to
transform processes across the broader digital enterprise and to leverage this common view of the
customer and the strategic value of data in coordinated real time systems of prediction and decision-making.

Front office vendors have an abundance of market opportunity to replace existing systems for enterprises wishing to enhance their customer facing interactions and transform their businesses. Over time, there exists the opportunity to leverage their growing market positions to expand into HCM, next generation financials and broader real time ERP. Today, this is likely to be viewed as a distraction, so the markets would likely need to mature more before we would expect the front office vendors to consolidate into the back office.

The ERP vendors have a lot of work to do to re-platform, but may have a greater desire to consolidate further into the strategic front office markets to deliver a more comprehensive, integrated systems and expand their foothold in the front office as a driver of digital transformation across industries.

Integrating front office and back office systems is technically challenging given they were developed separately. Also, valuations of front office vendors is high, so economically this may be challenging as well until market growth opportunities mature. In theory, the market opportunity exists to leverage the new platform for a more integrated real time enterprise solution set that more fully leverages the new strategic value of data and algorithms as a competitive differentiator and a defensive moat. Perhaps the public cloud vendors are the more likely consolidators. The challenge here for the public cloud vendors would be the change in business model associated with moving into the enterprise applications market and the significant change in distribution and support required with direct sales and support organizations.

We would also expect the need to transform business processes broadly in the market, not just by reengineering existing ones, but by enabling a new generation of processes that were not possible before the advent of these new predictive and real-time capabilities. We foresee a renaissance of business process reengineering and business process management (BPM), with a need for low- and no-code tools given a likely degree of magnitude increase in new business processes likely to emerge on the new platform. Big data databases, cloud native data and application integration tools and cloud-native middleware likely also would be more in demand.

**Data is the new strategic competitive advantage** as companies across all industries attempt to leverage data and software to transform their businesses. Across all industries, digital change must not be incremental. Rather all enterprises must fundamentally transform how they interact with customers, operate more efficiently and in real time, and optimize production and logistics.

Companies’ products must be digitally enabled to provide direct connectivity to the consumer given the critical role that data plays in enabling the digital enterprise. Digital transformation and the accelerated adoption resulting from a new generation of predictive systems will also alter business models across industries and enable new services and revenue streams. Data is critical to every enterprise going forward.

Traditional economics define inputs to GDP to include land, labor and capital. We will now need to add to this list the strategic value of digital data, which every company will need in order to adapt, be competitive and enhance productivity. Data is the critical enabling differentiator. It feeds the machine-learning engines that produce predictive capabilities to facilitate transformation of business processes and drive new business models and enhance productivity across economies.

A direct relationship with the end consumer is often key to acquiring and fully leveraging data for product enhancement and expansion of new services as part of the digital transformation and resulting changes in business models. The recurring revenue streams associated with service-enabling products is potentially high incremental margin revenues.

The move to cloud computing and an increasingly digital customer experience have already spawned the digital transformation of industries, but this is just the beginning. It is the foundation for the next wave of innovations that will empower a new generation of predictive applications and a renaissance of business process automation to accelerate the transformation of industries. We take a closer look at the underpinnings of this platform shift and the historical context for this in Figure 5.
Commercial computing has evolved over the past 60 years, starting from its origins on centralized mainframes in the back office to automate previously manual and paper-based record-keeping functions and workflows. Subsequent generations have largely delivered better ways to automate previously manual processes and workflows, along with some innovation in how companies engage with their customers. The next wave, as we discuss, will be about automating processes that did not exist before, because we did not have the big data sets to drive AI and machine learning for predictive capabilities and real time enterprise systems for automated decision making.

The centralized, mainframe era was followed by a second platform of computing, which began about 25 years ago, benefiting from advances in and the superior price/performance of microprocessor-based systems and pervasive networking. This enabled distributed client server computing and a proliferation of departmental applications and front-office automation. It was accompanied by a wave of business process re-engineering, which began in the 1990s as businesses adopted new and standardized industry best practices and processes facilitated by distributed computing capabilities.

With subsequent advances over the past 10 - 20 years in cloud computing, together with ubiquitous broadband internet and, more recently, smartphones/mobile computing, we have seen a third shift of the computing platform to computing-as-a-service (SaaS, IaaS and PaaS). Networking, compute and storage are delivered as a service by public cloud vendors (Amazon Web Services, Microsoft Azure, Google Cloud Platform and others) and are disrupting the previous generation of “on-premise based” hardware, storage, networking and software vendors. Companies no longer need to build their own data centers. Instead compute resources can be consumed, and charged for, based on a consumption model similar to the public utilities model. This has shifted IT from more of a capital cost to an operating cost model. As Charles Phillips (CEO of Infor) stated at an AWS Re:Invent Conference, “Friends no longer allow friends to build data centers.”

The fourth new platform of computing emerging leverages the superior scalability and economics of storage and compute offered by the public cloud computing vendors together with a number of additional enabling new technologies. The convergence of these additional enabling technologies notably includes the introduction of big data databases hosted economically in the cloud. It also includes the ability to: build big data sets that previously would not be feasible; apply machine learning to make sense of the big, complex data sets of the increasingly digital enterprise; and build predictive algorithms that can enable a new generation of intelligent or predictive applications. The pervasive use of smartphones and mobile apps as a consumer interaction platform and low-cost sensors for IoT are also complementary enabling components of the new platform (shown in Figure 6).
Closer Look at the Emerging Fourth Platform of Computing

Each of the tech industry’s platform shifts has been driven by a synergistic convergence of enabling new technologies. These shifts in computing architecture have become a familiar cycle in technology (Figure 6).

The emerging Fourth Platform of Computing is being enabled by:

1) The enormous capacity and low cost of networking, compute and storage in the cloud (from Amazon, Microsoft, Google and others).

2) The open-source foundation blocks of the new technology stack (Apache Hadoop infrastructure stack).

3) Smartphones as a mobile platform with robust app ecosystems, geo-location technology, video, social media, ubiquitous broadband internet connectivity and, increasingly, virtual and augmented reality.

4) Big data architectures that enable users to begin to build big data sets that were not possible or economical in the past.

5) The proliferation of low-cost sensors to generate the digital telemetry that fuels some of the big data needs of this architecture (notably in IoT).

6) The ability to use big data sets to train ML engines, the output of which, in an operationalized capacity, generates real-time and predictive capabilities.

7) The availability and democratization of AI and ML as a service in public cloud platforms, which are inexpensive and available as part of the cloud infrastructure stack.

8) A secure and pervasive wireless and mobile network for connectivity and delivery of the high volume, velocity and variety of data, which fuels the next-generation platform.

FIGURE 6: NEXT-GENERATION COMPUTING PLATFORM

Platform shift to cloud native, scale-out architecture to enable ML and predictive applications – likely a disruptive shift in software

Fourth IT Wave Is Under Way

Enabling the Next-Generation Computing Platform, IoT and Digital Transformation
The digital transformation of industries is of great strategic interest as company boards and management teams worry about being Amazon’d, facilitated by the growth and record-level VC funding for opportunistic and forward-thinking disruptors in every industry. Some companies start out as digital first companies while others must transform themselves.

Technology innovation builds on the convergence of new technologies that leverage prior innovations. It is the convergence of these new technologies—which can combine in a synergistic way to provide an enabling new platform—that would have a more profound impact than any one component in isolation.

For example, Uber has leveraged a number of these complementary new components, including consumer smartphone/apps technology to disrupt a market. To establish a community of service providers and customers in building an efficient market, mapping and geolocation/GPS were essential elements for location information—likewise, Waze revolutionized how we drive given the smart use of traffic data for route optimization and drive times—e-payment and on-line credit card payment for seamless real-time billing was essential, as were online support from a mobile smartphone app and a rating system for driver and consumer feedback to influence behavior in the marketplace.

Only several years earlier, the enabling components may not have been mature enough or widely adopted to facilitate commercial success of these new business processes and disruptive business models. It was the synergistic convergence of complementary new technologies that enabled the next generation of applications and disruptive new approaches and business models. We now have a very rich new platform of enabling technologies for innovation, particularly when we consider the application of machine learning to big data sets for predictive capabilities.

We are in the early innings of the next big platform shift in computing architectures. It is the application of AI to large data sets that is architecturally challenging for traditional systems as enterprises move/transform to e-commerce and other data-driven businesses. It’s not only about the data that we collect for successful transactions, we have not yet begun to look more broadly at the data for transactions that did not happen to better understand why these customer experiences failed. The size of the data sets going forward will be enormous and enable much greater predictive capabilities. The competitive advantage of data that can be used to create better and better algorithms (the “secret sauce,” or IP, of the digital enterprise) will drive us to non-linear growth in demand for more and more data, storage and processing power.

We have been watchful for companies that are starting out on the cloud platforms and leveraging the AI and machine learning delivered as a service in the cloud stacks to deliver predictive applications. We are now beginning to see a few emerge, but it is still very early in the platform shift.

Renaissance of Business Process Reengineering

Every application offered by the apps vendors will need to change, to varying degrees, to embrace or accommodate the predictive capabilities enabled by the new platform, and at a more granular level there will be processes to be automated that did not exist before. Traditional applications automated previously manual processes and mirrored earlier paper-based workflows, but the new platform makes possible business processes that did not exist before, enabled by the real-time and predictive nature of the new technology stack.

The innovation of business process re-engineering we saw in the 1990s, enabled by distributed client server technology, is likely to be repeated, this time facilitated by the emerging new technology stack and a new generation of predictive applications. The digital transformation of business leverages this new technology stack and embraces a shift in strategy in how the enterprise business functions in the future as a digital enterprise (the new business processes to be automated) and implications for how the business model is impacted.
For example, one of the most pressing changes is to the entire customer interaction experience to enable smartphone apps and mobile assess for a broader, omnichannel seamless experience across sales, service, support and billing interaction. (Omnichannel marketing refers to the multi-channel sales approach across mobile, on-line, social, phone or traditional brick-and-mortar store fronts with a seamless experience across all.)

Organizations can be reengineered to leverage the capabilities of the new platform to deliver real-time personalization and recommendations to shoppers and to capture the changing demand pattern and the impact on forecasts. In a real time enterprise, this data flows through to update production planning and logistics and is looped back to inform the marketing process with constantly updated data.

We have focused here primarily on the implications of the platform shift to the enterprise software and broader tech sector, but this new platform also cuts horizontally across all industries, further enabling digital transformation and the potential for business model disruption and opportunities for market expansion in health care, retail, aerospace, media & entertainment, financial services, transportation, energy, etc. A comprehensive discussion of digital transformation across all industries is beyond the scope of this discussion, so we have touched on this in summary fashion in Figure 7.

The industrial sector has been among the first to adopt this new platform, so we will discuss briefly below how the platform has evolved as a predictive system.

**FIGURE 7: CRITICAL USE CASES ACROSS MANY INDUSTRY VERTICALS**

Critical Use Cases Across Many Industry Verticals

*Impact of Machine Learning, IoT and Digital Transformation*

Source: Perella Weinberg Partners
The New Computing Platform Operationalized in IoT

Internet of Things (IoT)—new computing platform and predictive apps operationalized for industrial/manufacturing, energy and other markets

The new computing stack: The application of the new computing infrastructure stack for IoT-related use cases, as shown in Figure 8, has been operationalized by innovative companies, including Tom Siebel’s C3 IoT, PTC’s Thingworx, GE Predix, and evangelized in two definitive articles on IoT published in the Harvard Business Review, co-authored by Professor Michael Porter and PTC CEO Jim Heppelmann (Figure 9). IoT is the application of the new generation of predictive systems and applications to digitally transform production processes, broadly defined.

Gartner now references nearly 400 IoT platform vendors; although we note that there are few complete, integrated platforms. Most vendors address only a few layers of the stack such as aggregating sensor telemetry and many are in industry- or company-specific silos. IoT stacks are an early example of the new generation, cloud-native computing stacks. The public-cloud vendors are beginning to offer their own open-source IoT stacks, complemented in partnerships with other vendors to leverage the industry-specific domain expertise and applications for industry verticals.

The ability to ingest massive amounts of real-time data is at the core of the stack. This could be sensor telemetry from a jet engine or from a smart meter in the energy field. There is a need for a big data database to collect the data for the initial phase of training the ML engines, but subsequently the data layer needs to be less of an Hadoop data lake big data store and more of a Cassandra-like time series store, or other high-performance NoSQL database such as MongoDB, or converged platform such as MapR that offers high performance in aggregating the high volume, velocity and variety of data to feed the ML algorithms for predictive outcomes. Relational databases store data in rows and columns, which is optimal for transaction processing but not for unstructured big data.

The ability to operationalize the ML algorithms and to deliver highly scalable performance with the platform, across millions of devices, all generating real-time telemetry is more difficult than simply assembling a collection of open-source Apache Hadoop components. Proven scalability is key, as is the comprehensiveness of the platform, with many IoT offerings only addressing a portion of the stack. Collecting sensor data and aggregating this into a data store does not compare with the functionality of a comprehensive stack, as shown in Figure 8.

The platform also needs to offer integration with legacy systems. If the ML layer predicts that an asset is about to fail, there is a need to gather data about that asset, which is likely contained in the SAP or other PLM (product lifecycle management) system, which might help in the assessment of whether the assets should be repaired or replaced, and help inform the service department of what action is to be performed. Application layer integration is an essential component of a complete IoT stack, given that much of the data is locked up in traditional systems of record.

An example of a next-generation predictive application would be to leverage this predictive data of equipment failure (derived from the ML algorithm) to inform the CRM applications layer, instructing the sales and support departments to proactively initiate sales and support contact with the customer. Based on the asset data coming from the PLM system, the recommendation may be to replace versus repair, for example, a part. Also the services module is also informed in real time of the need to schedule a preventative maintenance visit, and since this is integrated with the PLM data, parts can be proactively specified and provisioned for the field service team.
The CRM system becomes more of a lead-generation system, and seamlessly integrates with the field service system, which is also informed in real time of the need for service and automated scheduling as well as informing customer support with relevant information. This should also inform the production and engineering systems of potential product issues. Also, the data in a real-time enterprise should flow through and integrate with billing, inventory, production scheduling, etc.

Predictive capabilities apply to all applications, enhancing recruiting based on automating resume scanning or correlating performance with candidate skills, marketing automation with the enormous volume and velocity of real-time sales and customer interaction data, optimizing production and logistics by better understanding large volumes and varieties of data. Budgeting and corporate performance monitoring would also be an opportunity for predictive capabilities along with security monitoring and anomaly detection.

**Companies today are beginning to collect proprietary big data sets for the purpose of training the ML engines.** Much of the data that might also need to be considered may be in public or third-party data sources, such as weather, traffic, GIS (Geographic Information System) data etc. The flexibility to consider a wide range of data sources is a key advantage of the extensibility of the new platform versus existing scale-up architectures based on relational database technology.

**Identity and security** are, of course, also important parts of the new technology platform. With an estimated 8.4 billion connected devices growing to a projected 20-30 billion by 2020, the expanded attack surface for exploitation by cyber criminals or hostile nation states is enormous. Devices such as lighting systems, smart TVs, intelligent HVAC systems, security systems, jet engines, connected vehicles may represent a rogue network that is not being managed by traditional IT, but instead providing a door into the enterprise-wide networks.

---

### IoT TECHNOLOGY STACK

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tools that manage user authentication and system access, as well as secure the product connectivity and product cloud layers</td>
<td>Application Development Platform</td>
<td>A gateway for information from external sources - such as weather, traffic, commodity and energy prices, social media and geo-mapping - that informs product capabilities</td>
<td>Tools that integrate data from smart connected products with core enterprise business systems such as ERP, CRM and PLM</td>
</tr>
<tr>
<td>Real-Time Machine Learning / Analytics</td>
<td>Data Aggregation (Hadoop or NoSQL)</td>
<td>Real-Time Streaming (Spark)</td>
<td>Network Connectivity</td>
</tr>
<tr>
<td>Edge Processing</td>
<td>Protocols</td>
<td>Sensor Telemetry</td>
<td>Sensors</td>
</tr>
<tr>
<td>Devices</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### COMMENTARY

- An IoT platform (on-premise or cloud hosted):
  - Collects real-time streams of sensor telemetry
  - Provides for big data storage (Hadoop or NoSQL)
  - Applies machine learning and analytics to the real-time data streams to make sense of the high volume, velocity and variety of live data
  - Provides for integration with legacy systems of record
  - Provides an application development platform to automate the development of a new generation of business process automation

**Strengths**

- High level of interest in IoT is translating into proof-of-concept and pilot IoT projects for many enterprises
- Enterprises are starting to use IoT-enabled data streams to provide new and improved services and offerings to consumers and businesses

**Limitations**

- Few vendors today have a complete IoT platform
- Landscape is highly fragmented - partial technology stacks and hundreds of companies in the IoT platform space
- Nascent market with vendors focused on specific industry verticals

---

Source: Perella Weinberg Partners
The Mirai attack in October 2016 exploited the millions of IoT devices that were remotely compromised, and infected companies globally, including Amazon and Netflix, with massive distribute denial of service (DDOS) attacks. Traditional endpoint protection does not apply if the devices were not designed to host an agent and receive updates and cannot be managed remotely. The convergence of operations technology (OT) with information technology (IT) underscores the need to apply IT to a class of devices that maybe did not anticipate this requirement. Security is clearly an important element of the new stack and a growing market opportunity.

**At the top of the new architecture or infrastructure stack is the application development facility** (Figure 8) or tools to help build this new generation of applications that are informed by the predictive capabilities of the ML algorithms.

Traditionally, applications and computing have been used to automate previously manual processes (record keeping, accounting, payroll, managing contacts and relationships); later we saw systems of engagement in automating how one addresses the customer, from marketing to sales to customer support. In the next generation, we will be able to automate processes that have no historical manual analog—something not possible before.

For example, the ability to predict when equipment is likely to fail (Asset Performance Management is a killer app for IoT in the industrial sector) can inform the CRM system and the previously static sales automation system can become a lead-generation system, proactively alerting the sales organization that the customer will soon need to repair or replace a piece of equipment as noted earlier. The sales and customer support functions are significantly enhanced with the scale-out, big data architecture and real-time predictive capabilities as output from the machine learning layer. This workflow or process did not exist previously and will now need to be automated. Estimating the life of an asset can now be a new and better informed process, not possible before the application of big data. There is a need to design new processes and workflows to be automated with low- or no-code tools.

We estimate there could be a degree of magnitude increase in the number of new processes that have not been automated previously and the demand to automate these processes along with the overall digital transformation of all industries and government functions will result in a renaissance of business process reengineering. Consulting and computing services firms will need to hire thousands of data scientists and consultants and assist their clients digital transformation journey begin with a comprehensive digital strategy. BPM (business process management) will likely be reinvigorated for the professional services firms.

Tesla uses hundreds of sensors in its vehicle and the Airbus A380 has 70,000 sensors. The growth in sensors for IoT and big data analysis will require business process automation as organizations need to address what is likely to be a significant increase in new processes/applications resulting from the adoption of the new computing stack. We expect every application in the ERP suite will be enhanced by this new platform, but the bigger opportunity may be in the expansion of the market for new processes and applications that have not been possible until the digital transformation of industries, enabled by the new technology stack.

Much of the data that will be acted upon will come from the edge of the networks. This could be sensor devices or processors such as the AI processors being developed for deployment in autonomous vehicles. The ML algorithms may be developed in the cloud, but deployment may be at the edge. An autonomous vehicle may need to ingest 2 GB of data per second and latency is not an option, so deploying the ML algorithms on the edge device is required.
Data science is being applied to analytics and ML and the models are being integrated and operationalized to make sense of and act on the enormous volumes and velocity of data coming in at real time. Products themselves are being redesigned to accommodate connectivity using low-cost sensors, creating a market opportunity for adaptive systems and a new generation of real-time, smart and connected products.

**New stack enables the real time enterprise:** The new technology stack will be the nervous system that connects and enables collaboration among previously separate business functions, with greater interoperability across product development, manufacturing, marketing, sales, service, support, finance and human capital management. These systems of prediction will have the ability to loop back and control the OT assets being monitored to optimize production. Continuously updated demand data along with real-time changes in production can be reflected in dynamic supply chain planning and management of new business processes.

An important element of the connected enterprise will be the enablement of real-time capabilities and automated decision-making. The enterprise applications market will likely undergo transformative growth, as we saw with the earlier generation of business process reengineering as the market moved from mainframes to distributed client server technology. The catalyst for growth in applications will be the enablement of real-time, predictive capabilities provided by the new supporting infrastructure stack.

We hope to stimulate further discussion on what appears to be the biggest platform shift in computing history as it evolves. We welcome your thoughts and suggestions. We would expect to offer thoughts in more detail on IoT and more broadly on the impact of digital transformation across industries.

---

**RICK SHERLUND**
Partner, Co-Head of Technology
(415) 671-4557
rsherlund@pwpartners.com

Rick is a Partner at Perella Weinberg Partners in the Advisory business and is located in San Francisco. He is a leading authority in the computer software industry. He is driving the Firm’s coverage of software clients, as well as developing a broader technology franchise with his Advisory Partners to assist clients across all industries as digital transformation redefines their businesses.

His background includes 25 years at Goldman Sachs where he was a Partner Managing Director, Head of the Technology Research Group and a leading software industry analyst - ranked as the #1 software industry analyst by Institutional Investor for 17 consecutive years. He was involved in the initial public offerings of many of the industry’s leading companies and has long been a strategic adviser to companies positioning themselves for upcoming new platforms of computing. Over the past several years Rick has focused on the digital transformation of enterprises, IoT and machine learning and the upcoming platform shift from systems of record and systems of engagement to a new generation he characterizes as systems of prediction.
Legal Disclaimer

This Presentation has been provided to you by Perella Weinberg Partners and its affiliates (collectively “Perella Weinberg Partners” or the “Firm”) and may not be used or relied upon for any purpose without the written consent of Perella Weinberg Partners. Reproduction of this Information, in whole or in part, is prohibited. These contents are proprietary and a product of Perella Weinberg Partners. The Information contained herein is not an offer to buy or sell or a solicitation of an offer to buy or sell any corporate advisory services or security or to participate in any corporate advisory services or trading strategy. Any decision regarding corporate advisory services or to invest in the investments described herein should be made after, as applicable, reviewing such definitive offering memorandum, conducting such investigations as you deem necessary and consulting the investor’s own investment, legal, accounting and tax advisors in order to make an independent determination of the suitability and consequences of an investment or service.

The information used in preparing these materials may have been obtained from or through you or your representatives or from public sources. Perella Weinberg Partners assumes no responsibility for independent verification of such information and has relied on such information being complete and accurate in all material respects. To the extent such information includes estimates and/or forecasts of future financial performance (including estimates of potential cost savings and synergies) prepared by or reviewed or discussed with the managements of your company and/or other potential transaction participants or obtained from public sources, we have assumed that such estimates and forecasts have been reasonably prepared on bases reflecting the best currently available estimates and judgments of such managements (or, with respect to estimates and forecasts obtained from public sources, represent reasonable estimates). The Firm has no obligation (express or implied) to update any or all of the Information or to advise you of any changes; nor do we make any express or implied warranties or representations as to the completeness or accuracy or accept responsibility for errors.

Nothing contained herein should be construed as tax, accounting or legal advice. You (and each of your employees, representatives or other agents) may disclose to any and all persons, without limitation of any kind, the tax treatment and tax structure of the transactions contemplated by these materials and all materials of any kind (including opinions or other tax analyses) that are provided to you relating to such tax treatment and structure. For this purpose, the tax treatment of a transaction is the purported or claimed U.S. federal income tax treatment of the transaction and the tax structure of a transaction is any fact that may be relevant to understanding the purported or claimed U.S. federal income tax treatment of the transaction.