



Perspective

As 2019 begins, the Big Data landscape continues to mature:

- The product ecosystem is starting to consolidate around a small number of core technologies such as Parquet, Spark, Cassandra and Kafka. Newer players may have better ideas, but are finding it difficult to get funded or build traction with customers.
- It is far easier to identify use cases where Big Data provides a clear ROI, based on the experience of early adopters. For example, in the financial sector, Know Your Customer (KYC) is a ubiquitous regulatory requirement that requires analyzing data from all parts of the business in order to reconcile them into a single consistent picture of the customer. For retail, the use case with the most visible ROI is personalization e.g., adjusting the web landing page so that the customer sees items that are directly relevant to his/her interests, or creating personalized coupons and email campaigns.
- The supply of skilled Big Data manpower, while still far short of the demand, is being bolstered by a new generation of college graduates for whom Big Data has become a standard part of software education.

At the same time, there are a number of new developments and trends that could have an impact on your Big Data plans:

- Internet of Things (IOT), which has been unfairly maligned as a trend that never happened, continues its steady growth, in areas like smart cars and smart cities. IoT will contribute to massive growth in the amount of data that needs to be analyzed.
- Cloud providers like Amazon and Microsoft continue to innovate on Big Data products in order to “lock in” their customers. The reasons to use a proprietary product rather than an Open Source alternative will continue to grow.
- Serverless Big Data databases, which deploy computing resources on demand, are entering the mainstream, because they can be an order of magnitude less expensive than traditional, dedicated resource databases in the Cloud.
- Data ingestion tools are emerging that can dramatically shorten the time required to cleanse data sources (currently 60- 80% of any Big Data project) by introducing more Artificial Intelligence into their products.
- Self-service data analytics tools continue to flourish. These domain-specific tools enable non-data scientists to explore their data and arrive at insights via easy-to-understand data visualization.
- GDPR regulations, which take effect in 2019, will require many companies to redefine their data infrastructure, e.g., to separate out sensitive data that can identify a customer into separate databases.
- Data security continues to be an issue, as hackers continue to overrun the defenses of outnumbered enterprise data security officers.



Next Steps

If you are looking to launch a Big Data initiative, you probably do not have the luxury of waiting before starting your project. In most business sectors, a large competitor has already declared success based on insights derived from some Big Data initiative, and the rest of the industry has to react now or risk being left behind.

Whether you are just getting started, or have already launched a Big Data initiative and are not seeing the successful results your competitors are touting, you need to take advantage of the experience of others. The field is full of competing or overlapping products, each of which claims to be the right solution for Big Data. Left to evaluate this cacophony of conflicting voices is your organization. It's hardly a fair battle –the only way to cut through the hype around a product is to try it yourself, and/or talk to someone you trust who is using it.

The best advice to a company contemplating a Big Data initiative or seeking to right the ship on an existing Big Data initiative is: Get help. Perhaps you can hire experts from outside your organization who have already built such a system. Perhaps you can build up a network of friends who have made the beginner's mistakes and can share their insights with you. Another option is to partner with a company like Ness Digital Engineering that has managed a broad range of Big Data projects and technologies, and has a proven track record of success.