

Technology	What it is, Why it Matters
<b>Hadoop</b>	Perhaps the most-often cited technology in connection with Big Data, Apache Hadoop is an open source framework that allows for the distributed processing of large data sets across clusters of inexpensive computers. Often used to analyze customer behavior at large retailers, it could also be used to analyze security events or process large image sets to look for meaningful changes over time. Hadoop is at the center of a cluster of open source projects, which include tools for management, data collection, storage and machine learning, to cite a few examples.
<b>MapReduce</b>	Usually mentioned in the same breath as Hadoop, MapReduce is a programming technique for processing parallelizable problems across very large data sets. Many aspects of public safety systems would fit into this framework as large networks of sensors could be individually analyzed in parallel to one another to extract information and patterns that could then be summarized at a higher level.
<b>NoSQL databases</b>	NoSQL is a class of database systems that can be used when the underlying data relationships are not usefully understood through the traditional relational database model. The emphasis of these databases is on storing and retrieving very large data sets for real-time or statistical analysis. An example might be the large stream of motion detection and other sensor events that come from an enterprise or municipal surveillance system.
<b>Columnar databases</b>	A type of database in which storage access is optimized around columns rather than rows, which is more efficient for certain type of computations, particularly those based on aggregation of columnar data sets. This technology can reduce storage costs and speed data access for certain classes of problems, possibly edging proposed solutions closer to feasibility or ROI.
<b>In-memory databases</b>	In-memory databases are those that fit large data sets into main RAM memory as opposed to disk file storage. This approach can speed query time by orders of magnitude – a make-or-break difference. In-memory databases have become practical, as RAM prices have fallen dramatically over the past decade. They are useful when response-time is critical, which again includes many life safety and loss prevention scenarios.
<b>Advanced storage</b>	None of these technologies would have gained a foothold were it not for the availability of very fast, inexpensive storage technology. While disk drive prices have continued to drop, solid state drives have further reduced access times for primary storage, and both can be joined seamlessly with the virtual storage management software tools that are needed for very large distributed data sets. Most large enterprises are already using these storage technologies for their own corporate data, and would recognize the value in reducing loss or minimizing risk through large-scale security event processing.
<b>Talent</b>	Talent has been added to this list because without the right type of data scientists in an organization, none of the technologies will yield any useful results. This is a big shift for most organizations, and the physical security world in particular, as deep analytics has not been a major focus of our products, outside of national security applications, classified or otherwise. In this regard, McKinsey predicts that by 2018 there will be a shortage of 140,000 to 190,000 people for these needed skills.