Big Data and Privacy for Physical Security

By Ray Bernard, PSP, CHS-III

Version 1.0

Big Data and Privacy for Physical Security

By Ray Bernard, PSP, CHS-III Version 1.0

Big Data and Privacy

Big data and *privacy* are two interrelated subjects that have not warranted much attention in physical security, until now. Both subjects are about to become of strategic importance to security, due to recent advancements in video analytics and big data technologies, court rulings regarding data privacy rights relating to surveillance video, and the growing value of operational data that can now be extracted from video surveillance systems.

On the one hand, increasing concerns about privacy rights — and the corporate liability for neglecting them — are already having chilling effects on thoughts about sharing security video and video-derived data beyond a limited set of security personnel. On the other hand, utilizing security video for business intelligence purposes is now a reality in retail point-of-sale and dozens of other applications. Video information sharing provides a new high business ROI for the significant investment in security video, and makes a compelling case for pursuing its non-security business applications.

There is no doubt that security practitioners and their service providers will be called upon to resolve this dilemma in a way that results in maximum benefit to their organization, without compromising any privacy-protection obligations.

Advancing Information Technology

Central to this picture is advancing technology, including cloud computing, which is driving massive change in the ways that people, businesses and governments interact, as well as the way that organizations operate. It also impacts the nature of any IT department's work. The collaboration between security departments and IT departments will become much more important than in the past due to:

- a greater use of wired and wireless networks for security systems, including corporate networks;
- the increasing business value of information from video systems and video analytics;
- the growing expectation that organizations will utilize mobile device applications to support emergency response, crisis management, and individual safety and security on company premises; and
- the growing relevance of big data and its technology to security operations.

This is the organizational and technological context for a security department's use of big data and the consequential need to address privacy issues.

This paper's purpose is to provide an overall understanding of emerging big data technology, and of the business expectations regarding big data. The paper also provides guidance on how privacy obligations and confidentiality concerns can be effectively addressed when

using security systems and devices that track the behavior of individuals or contain sensitive information of any kind. At several points in this paper, retail store video analytics are used as an example of a current-day physical security application of big data technology. It is a good example because the use of video analytics has all the technical and organizational dynamics that are typical of big data initiatives.

The Transition to Information Technology

It is the arrival and continuing advancement of information technology that has brought us to the point where big data is relevant to electronic physical security systems, and where the confidentiality of personal and organizational information is a growing fundamental concern for physical security applications. From the 1960s to the 2000s, security systems technology began evolving from manually controlled electric and electro-mechanical devices ("buzz" the door open); to solid-state electronic devices (settable schedules and alarm annunciation); to hard-wired systems (CCTV and card access control); to software-managed computer-based systems with databases (photo ID cards, computer display of alarms and video); to IP-networked systems and client/server software (security video DVRs for analog cameras).

The advent of the internet and World Wide Web ushered in IP video cameras and internet-hosted systems. Higher reliability and speed for networking and the internet made possible regional and enterprise-wide physical security systems. Today, cloud computing has enabled cloud-based security applications and security apps for mobile device users. The continuing miniaturization of electronics, low-power computing, wireless communication technologies, GPS and network location capabilities, and robotics technologies are revolutionizing video surveillance capabilities.

It is well past time to account for the fact that today's electronic security systems are highly connected IT systems, and that the standards and IT practices commonly applied to IT system deployments must be applied to electronic security system deployments as well, including data protection practices. Now that big data has arrived in the form of next-generation video analytics, security industry stakeholders (including end-user customers) need good answers to the following questions:

- How does big data technology work?
- What are the big data practices relevant to security applications?
- Why have privacy and confidentiality data safeguards become critically more important now?
- How has IT dealt with the fact that the three points above are still-evolving pictures?

The remainder of this paper addresses these and other related issues.

Part One: Big Data

Big Data Defined

Big data is more than just an extremely large volume of data, as explained in 2013 by ISACA (previously known as the Information Systems Audit and Control Association): "For some enterprises, big data is counted in hundreds of gigabytes; for others, it is in terabytes or even petabytes, with a frequent and rapid rate of growth and change (in some cases, almost in real time). In essence, big data refers to data sets that are too large or too fast-changing to be analyzed using traditional relational or multidimensional database techniques or commonly used software tools to capture, manage and process the data at a reasonable elapsed time."[1] What is considered to be big data varies, depending partly on the capabilities of the users and user companies and their tools. Furthermore, big data is a moving target because of the continuously expanding capabilities of the tools and systems for data processing, transmission and storage.

Big Data Definitions

The following are three useful definitions that reflect the expertise domains involved in utilizing big data.

big data

noun

1. *Data science.* Techniques and technologies for handling large sets of structured and unstructured data that have high *volume, velocity, variety* and *complexity* attributes. Related terms are *data mining* and *machine learning*.
2. *Data processing*. Data sets that are too large or too fast-changing to be analyzed using traditional relational or multidimensional database techniques or commonly used software tools to capture, manage and process the data at a reasonable elapsed time. (ISACA, 2013)
3. *Business intelligence, data stewardship.* The planned and documented extraction, distribution, utilization and protection of new and insightful information from verified large scale and/or or highly complex data sources, to inform business decisions.

For most medium and large size organizations, the security department's utilization of big data sources and technologies can be patterned after the successful practices of the IT department, especially those related to data governance. This is examined more closely in the section titled *Big Data Perspectives* on page 10.

1. ISACA.org, "Big Data: Impacts and Benefits," March 2013, <www.isaca.org/Knowledge-Center/Research/ResearchDeliverables/Pages/Big-Data-Impacts-and-Benefits.aspx>

Key Attributes of Big Data

The key attributes of big data were first described 15 years ago by IT industry analyst Doug Laney[2], referring to data sets that have three specific data processing challenges: *volume*, *velocity* and *variety*. Laney stated then, "Current business conditions and mediums are pushing traditional data management principles to their limits, giving rise to novel, more formalized approaches." As he predicted in his article, the volume, velocity and variety of data being generated has continued to explode, giving rise to dramatically different approaches in data management and data analysis than have been used in prior decades. Some big data experts add another attribute, *complexity*, because extracting and analyzing information from massively big data sets is a very complex process, and because there are likely to be very complex relationships among the various data elements.

Volume is the attribute most generally associated with big data. As of 2014, Facebook had an incoming daily data rate of about 600 TB, which they stored in a data warehouse sized then at 300 petabytes[3]. One petabyte (PB) is one thousand terabytes, or one million gigabytes. A Boeing 737 engine generates 10 terabytes of data every 30 minutes in flight[4]. As of December 31, 2015, Southwest Airlines operated 704 Boeing 737 aircraft, which average 11 hours and 10 minutes of flight time daily[5]. That means the Southwest fleet generates more than 326 PB of flight data daily — more than 544 times the amount of data that Facebook takes in daily. After a flight, Southwest Airlines will analyze the engine data and correlate it against weather conditions, flight occupancy rates, preventive maintenance operating thresholds and many other types of data.

During a flight, an information tool called Airplane Health Management (AHM), designed by Boeing and airline users, collects in-flight airplane information and relays it in real-time to maintenance personnel on the ground via a web portal. When an airplane arrives at the gate, maintenance crews can be ready with the parts and information to quickly make any necessary repairs.

Velocity refers to how fast the data is generated. The degree of challenge that data velocity poses is related to three factors: how quickly the results of the data analysis are needed; how complex and time-consuming the data analysis process is; and how costly the required computing, networking and data storage resources are.

Variety of data is a very significant complexity factor, as much of big data is not structured data in the way that traditional data is organized into databases. Traditional databases contain data records consisting of rows of individual data fields, with each data field containing a specific

2. Laney, Doug (2001); "3D Data Management: Controlling Data Volume, Velocity and Variety," gartner.com, retrieved 20 Aug 2016, <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

3. Vegata, Pamela and Wilfong, Kevin (2014); "Scaling the Facebook data warehouse to 300 PB," facebook.com, retrieved, <https://code.facebook.com/posts/229861827208629/scaling-the-facebook-data-warehouse-to-300-pb/>

4. Mathai, Paul (2011); "Big Data: Catalyzing Performance in Manufacturing," wipro.com, retrieved 24 Aug 2016, <http://www.wipro.com/documents/Big%20Data.pdf>

5. Southwest Airlines; "Southwest Corporate Fact Sheet", swamedia.com, 26 August 2016, <http://www.swamedia.com/channels/Corporate-Fact-Sheet/pages/corporate-fact-sheet>

type of data. For example, a contact record would contain data fields such as: Name, Company Name, Position Title, Email Address, Phone Number, Street Address, City, State and Zip Code. Structured data records such as these can be easily searched by Name, City, Zip Code, Company Name, and so on.

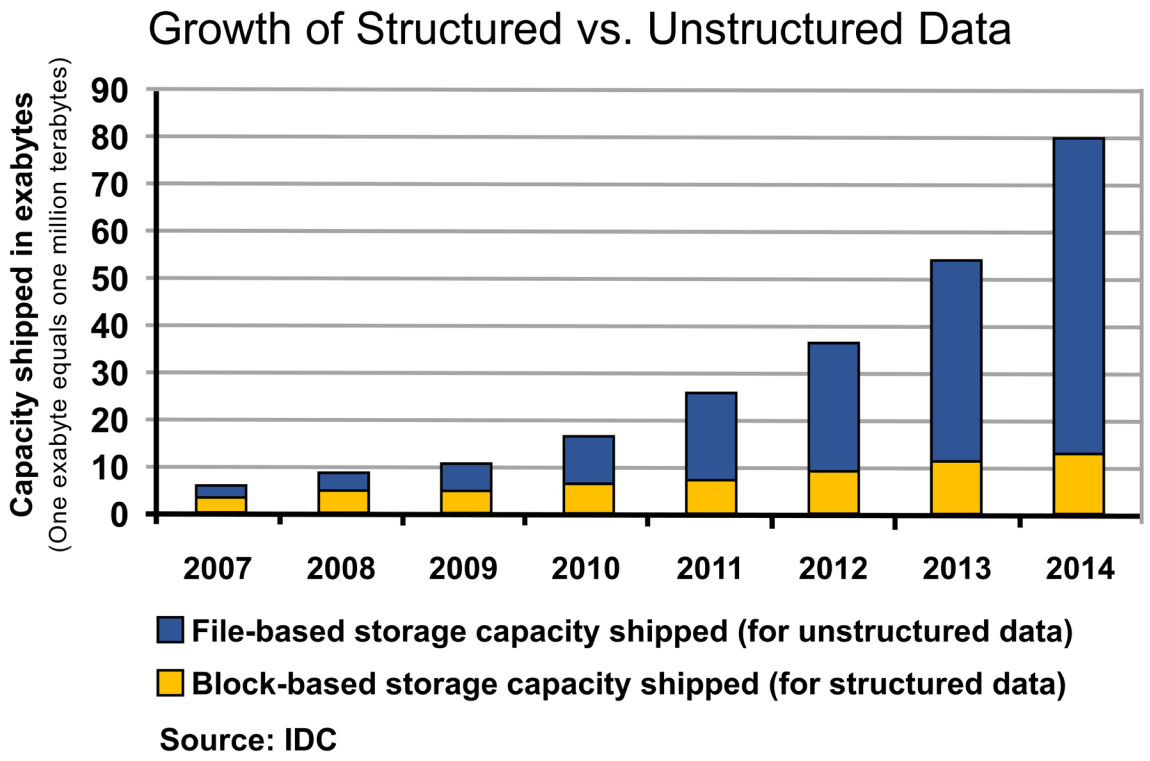
Boeing 737 engine data is structured data, while Facebook's data is mostly unstructured text data of highly varied content, plus pictures and their associated data including photo tags. In real time, Facebook does its best to use facial recognition analytics to identify the people that you know in photographs that you post, processing over 300 million uploaded photos per day[6]. Facebook and Southwest both deal with big data, but the volume, velocity and variety challenges they face are very different.

Forrester sums the whole picture up by defining big data as "Techniques and technologies that make handling data at extreme scale affordable."[7]

Unstructured Data

Big data is largely unstructured data, and the growth of unstructured data continues to outpace the growth of structured data, as shown in Figure 1 below.

Figure 1. Comparison of structured and unstructured data growth rates



6. Zephoria (Jul 2016); "The Top 20 Valuable Facebook Statistics – Updated July 2016," zephoria.com 25 Aug 2016, <https://zephoria.com/top-15-valuable-facebook-statistics/>

7. Forrester Research; "Expand Your Digital Horizon With Big Data," forrester.com. 30 Sep 2011, <https://www.forrester.com/report/Expand+Your+Digital+Horizon+With+Big+Data/-/E-RES60751>

Email messages are a mix of structured and unstructured data; a smaller amount of structured data and a larger amount of unstructured data. Sender, Receiver, Date, Time and Subject are a structured record. The body of the email contains unstructured data that is somewhat organized, and may or may not contain headings, paragraphs, bullet items or numbered lists. The body of an email may also contain small tables of data or pictures. An email message may or may not contain the sender's contact information at the end of the email body; when provided, it varies greatly from one sender to another. The email's content varies greatly from one email message to another. There is little consistency in the body content across thousands or millions of email messages. In corporate email, within a month's worth of emails there may be many messages about the same topic, such as project-related email or email planning company events. Relating data from one email message to other relevant messages is a huge analysis task, but it would have to be done in order to collect useful information from the wealth of data that resides in corporate email.

In big data analytics, structured and unstructured data are often analyzed in combination, whereby data in the unstructured data sets are associated with data in the structured data sets to provide deeper insight into the structured data. Sometimes the combination of structured and unstructured data, as in email, is referred to as *hybrid data*.

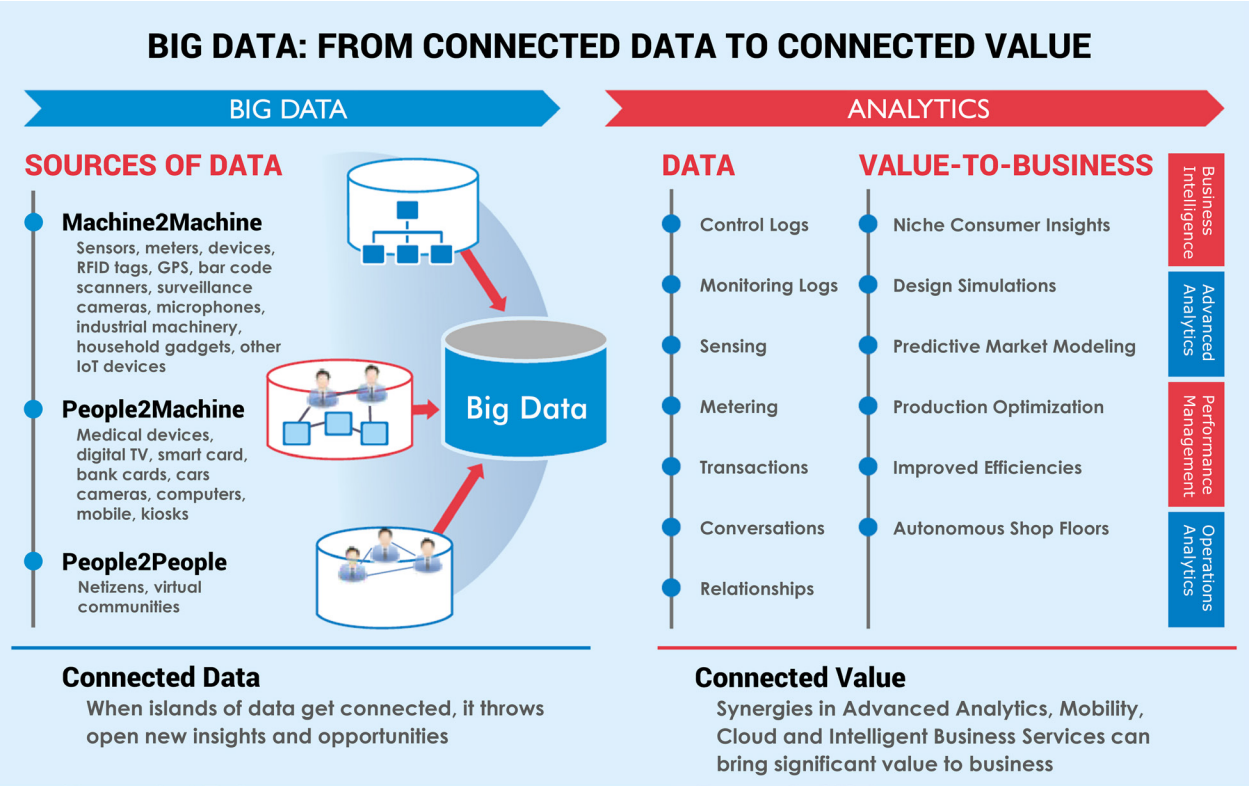
Surveillance Video Analytics is Big Data

Surveillance video analytics has all the attributes of big data. It has high *velocity.* For example, a Full HD surveillance camera transmitting 1080p MJPEG video at 30 frames per second will average 1.5 megabytes of data per second. A 100-camera HD video system has a high *volume* of data, as it can be collecting data at a rate of 150 megabytes per second, or much more for a system of high-resolution megapixel video cameras. The high volume is one of the reasons why the typical retention period for surveillance video is 30 days — it would take a sizeable data storage capability to maintain a year's worth of video stream data. Surveillance video data has high *variety*, as cameras capture activity of various kinds, and from one period of activity to another, no two video images from a camera will be identical.

Data Mining

The task of extracting meaningful information from massive amounts of structured or unstructured data is known as *data mining*. Data mining refers to extracting patterns and knowledge from large amounts of data, not extracting smaller parts of the data out of the larger mass of data. The value is in the meaning derived from the original data. Figure 2 (page 6) below provides an overview of big data value generation. *For a detailed look at the big data technologies used in advanced video analytics, see the SIA Technical White Paper, "The State of Security Video Analytics."*

Figure 2. Big data value discovery



Graphic concept and data provided by The Wipro Council for Industry Research

The initial value of structured data is the data itself. Its value is known before it is put into data storage, and it is stored in a structured fashion in order to facilitate accessing it specifically for its value. *Unstructured data contains potential value*, but the data has to be analyzed in order to identify and extract that value.

Different sets of structured data can be analyzed to create new information that doesn't reside in any of the analyzed data sets. This is known as *data correlation*, such as a correlation between the demand for a product and its price, which might be used to optimize pricing so as to maximize sales and profits. Data correlations are useful because they can indicate a predictive relationship that can support business decisions and business practice improvements. Data correlation is just one of many approaches to data analysis that are applied to big data.

Machine Leaning

Machine learning is the science of getting computer programs to modify and improve their performance through experiences, without having been explicitly programmed to make those improvements. Application programs that include machine learning can teach themselves to grow and change when exposed to new data.

The new generation of advanced video analytics includes machine learning functions. Machine learning can take place in any or all of three locations: video surveillance cameras, cloud-based

analytics systems and premises-based analytics systems. Video analytics machine learning involves building data models of video camera fields of view, including scene backgrounds, the objects in scenes (people, vehicles, bicycles, small animals, etc.), and their movements and activities. The data models are continuously updated, based upon ongoing analysis of scene objects and activities, and can even include object recognition (truck) and identification (FedEx delivery truck license plate 41033C1) as well as people recognition and identification (driver John Smith in FedEx delivery truck license plate 41033C1). Over time, machine learning improves the accuracy of the data extracted from video images.

Metadata

The data extracted from video images is called metadata (which means "data about data"). There are several types of metadata extracted for video analytics, as shown below in Figure 3 (page 9).

The previous generation of video analytics software did not store metadata. Each time a video analytic was run, the live or recorded video stream had to be reprocessed. Today's advanced video analytics store several types of metadata, which allow multiple analytics to be run at the same time off the original object and activity metadata extracted, using a fraction of the processing power required to extract data from the original video stream. At least the machine learning metadata must be preserved beyond the normal video retention period, so that it can quickly recognize objects and activity that cameras have seen in the past. Historical scene, object and activity metadata may also be retained longer to facilitate running activity and statistical reports; its data storage requirements are much less than that of the original video images.

As shown in Figure 3, there is machine learning metadata created and utilized in each stage of analysis: image object identification and labeling, object classification, object tracking, and object activity recognition. As the metadata (data model) for each analysis stage is improved, the results from one stage improve the analysis of the next stage, which in turn improves the following stage, and so on. The historical data about objects (people, vehicles, etc.) and their activities becomes more accurate over time.

Some of the video analytics metadata will include personally identifiable information (PII), which can be subject to privacy regulations and company policy that determines the data safeguards to be applied.

Figure 3. The generation of metadata in video analytics systems

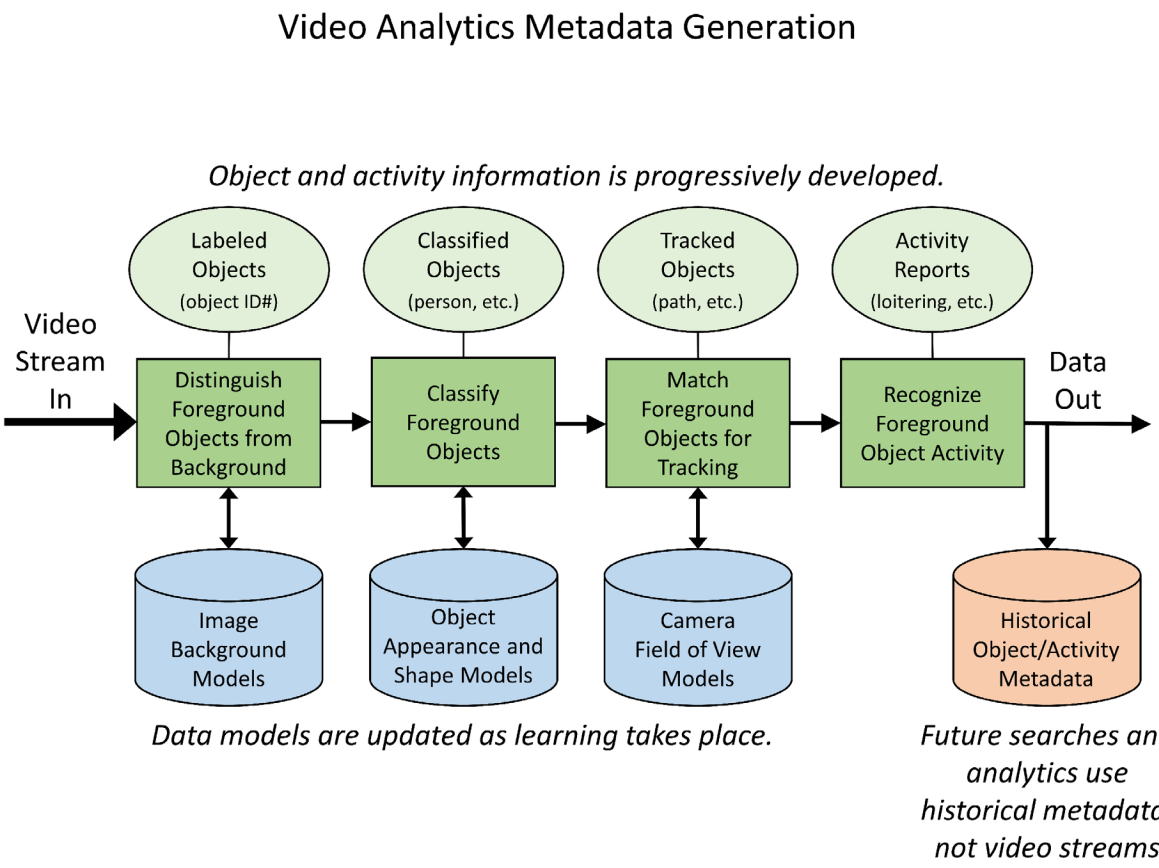


Image source: SIA Technical White Paper "The State of Security Video Analytics"

The machine learning metadata developed over time becomes more and more valuable, and thus the data stewardship roles are extremely important. The analytics metadata must be backed up and must be protected from unauthorized access. The machine learning metadata must be preserved well beyond the retention period for video recordings — especially when other business functions are making critical business decisions based upon the growing effectiveness of the business video analytics. It is a new situation that security departments can generate data that is critical to the decisions of other business functions. Organization policies will require that video analytics systems and all of their data receive the same safeguards and treatment as any other business-critical system, including periodic audits.

Big Data Perspectives

Most articles and papers about big data deal with one or more of the four business perspectives that relate to the use of big data:

- **Data:** what constitutes "big data", and where big data is to be found
- **Technology:** tools and applications used to manage or process big data
- **Business intelligence:** how big data value is created and utilized by an organization
- **Data stewardship:** protecting and managing data, data utilization and data processing performance

While the first three topics above have each received tremendous coverage on the web, relatively speaking, data stewardship has not. As with many roles relating to information technology, the concept of the *data steward* has evolved and expanded over time, partly because many more business functions are becoming data driven, and partly because so much more data is available now than has been in previous decades. This evolution of the scope and importance of data stewardship is one reason why the web contains such a wide variety of definitions and descriptions of data stewardship and the role of data steward. It is vital to get clarity on the topic of data stewardship, as the role is becoming increasingly more important since the advent and growth of big data.

Many IT security practitioners are very aware of the role of data stewardship. Big data has completely changed the types and amounts of information available for strategic and tactical decisions. This data growth makes data stewardship highly important, not just for IT but also for physical security, because big data growth has already started in the realm of physical security.

Data Stewardship

The overall concept of stewardship is the performance of responsible planning and management of resources. Data stewardship is the management and oversight of an organization's data assets to help provide business users with high-quality data that is easily accessible in a consistent manner[8]. Data stewardship ensures the integrity, usability, and security of the organization's data.

Data stewardship involves hands-on roles that are typically performed by a team or group of individuals, referred to as data stewards, each one focused on a particular part of the business, or on the technicalities of the data processes themselves. Data stewards apply their subject matter expertise in the business use of the data and/or the technologies that facilitate the data processing and handling, to achieve some part of the overall goal of data stewardship. One reason why the online definitions and descriptions of a data steward vary is that the various aspects of the overall data stewardship responsibility are split up among some number of data stewards.

8. TechTarget (Jun 2013); Definition: "data stewardship," techtarget.com, 20 Aug 2016, <http://searchdatamanagement.techtarget.com/definition/data-stewardship>

As the use of business insights derived from big data spreads throughout an organization, data stewardship becomes critically important for maximizing the value of the investment in big data resources, and for achieving the intended outcomes from big data initiatives. This is no less true for security departments and their use of video analytics for security and business purposes.

Data Stewardship Roles

Most data stewards do not hold full-time data management roles. Steward positions are often filled by subject matter experts that dedicate a portion of their time to data-related activities, while fulfilling a specific business or technical responsibility.[9] Data stewards' collective efforts improve the utilization and accuracy of data and information in a way that drives business performance and mitigates organizational risk.[10]

Data stewardship is part of the organization's *data governance* program. Data governance assures the availability, visibility[11], usability, integrity and security of the data employed in an enterprise. Data governance is a strategic function[12] that does not directly deal with data, but sees that the people, policies and processes are in place and functioning as they should be to manage the data assets. Data stewardship is tactical[13] in that each data steward's focus is on a particular set of data being used by or generated by a particular business function.

A large part of data stewardship is focused on answering the following questions, some of which have become much harder to answer with regard to big data:

What data do we need? In many cases this question has turned into an ongoing process, and is much more challenging to answer, as machine learning and artificial intelligence tools must be used to discover what kind of information is available in big data sets. If we don't know what information is available, how can we determine whether or not we need it? Even a small change in the way data sets are being evaluated can make a big change in the resulting information extracted.

How do we define it, document it and make it available? This involves identifying the key data, defining it and describing it, and figuring out how to store it and access it. Some decision points are, for example:

- How critical is the time factor for this category of information?
- Does it make a difference in security decision-making if the data is minutes, hours, days, weeks, months or years old?

9. Information Builders (2013), "Data Stewardship in Complex and Big Data Environments", informationbuilders.com, 20 Aug 2016, <http://www.informationbuilders.com/search/site/Data%20Stewardship?type=white_paper>

10. Ibid.

11. *Data visibility* refers to knowing what data is available where, and ensuring that business functions for whom the data has value are aware of it and can obtain appropriate access to it.

12. Haines, Rachel (May 3, 2012), "Data Governance is Strategic; Data Stewardship is Tactical," infocus.emc.com, 22 Aug 2016, <https://infocus.emc.com/rachel_haines/data-governance-is-strategic-data-stewardship-is-tactical/>

13. Ibid.

The answers will be different if the data is being used for a security operations center rather than a security risk assessment. Some countries may prohibit the use of specific data for certain purposes. Some data may not be allowed to leave the country in which it was created.

What are the business rules that should be established? These are rules that apply to the creation of the data, how the data is allowed to be used (a consideration that relates to data quality and data sources) and what other data may be derived from it (which can have privacy implications).

How do we document the data sources? With big data this can be more challenging than it may seem; such as for a situation where data analytics have provided valuable information derived from confidential or sensitive data sets. Can the derived data be de-classified if its value and content warrants it? If so, what can be said about the data's source?

What are the quality requirements for this data? This question is about the data being fit for the particular uses for which it will be applied, and it includes understanding what quality is possible or likely. It includes the objectives for using the data and the intended business improvement outcomes. The answer may impact earlier questions regarding who may be allowed to access the data and what business rules should be applied to it. For some data, answering this question may require an understanding of what kinds of further analysis will be applied to the data, and what quality is required to support such analysis soundly.

What should be done about any issues that arise around the use of this data? This can involve change management, as data user requirements can change, and what data is available can also change. It may involve data user education, instruction or training.

Who within the organization should know that this data is available, and how do we make it known to them? One stewardship responsibility is seeing that data assets are utilized to the maximum benefit of the organization, and that the value to the organization persists. Periodic reviews should be performed to consider the status of data utilization. Processes and procedures may be required to address when improvements in data utilization are needed.

What are the benefits of using this data, and how does that compare to the cost of obtaining and providing the data to its users? Since big data can involve multiple sources of data, and can require significant integration work for a variety of data sources, the integration cost factor cannot be overlooked. Answering this question will involve the subject matter expertise of data stewards who know the business functions where the data will be used, as well as technical data stewards who understand the data integration factors.

Additionally, data stewardship is responsible for carrying out the policies about data usage and security that have been developed through enterprise data governance initiatives. Data stewards provide liaison between the IT department and the business side of an organization. Even though it is a critically important function, data stewardship is often given less attention

than the other three big data perspectives. This has occurred primarily because the amount and scope of the data that is now available has not been available before, and because data stewardship requirements were simpler and less critical when the data stewardship role was first established. The importance of data stewardship continues to grow as the amount of data and the value of the data to the business continue to grow.

Physical Security Data Governance and Data Stewardship

Most physical security practitioners are very aware of their duty to protect sensitive security information, including surveillance video, and they develop appropriate policies and procedures to do so. However, security for sensitive information shared outside of the security department belongs under the overall umbrella of corporate data governance and its data stewardship activities. Because data stewardship involves more than information security, it behooves security practitioners to understand the full role of data stewardship as described above.

Take a case where video analytics is being used both for retail store security purposes and for store operations business purposes. For example, data analytics can be applied to staff behavior as well as individual and group customer behavior. There could be five data stewards: one each from security, store operations, merchandising, training and marketing.

The security data steward would oversee the configuration and continued operation of video analytics applications, and the safe handling of the video analytics data output. The previous sentence doesn't convey all the aspects of that responsibility, which are complicated due to continual advances in machine learning, evolving analytics technology, and evolving video metadata related to both security and business activities. *Both security and non-security data stewards would establish the analytics data objectives and data requirements for their business function's use of the data.*

The data stewards would determine the intended outcomes for their use of the analytics-generated data. Data stewards outside the security department would collaborate with the technical data stewards within the security department, and they would work out how they would access the data and what formats they would need for any data analysis tools; simple reports for human analysis may suffice, or data importable to spreadsheets or databases may be required. They would also work out what the data security requirements are, including for retention and destruction, and make sure that the appropriate personnel in their business function understand how to handle the data correctly. Sometimes this requires training and even testing, depending upon the sensitivity of the data and how widely the data is used within any particular business function.

Regardless of whether or not the label "data steward" is used, someone would have to perform the data steward functions described above. Doing so using a named role under the guidance of organizational data governance helps ensure that the organization receives maximum benefit from the investment in the video analytics program, and that its successes can be replicated elsewhere within the organization.

Previously, security systems have not generated data of such value to security departments and to organizational personnel outside of security. The proper care and safeguarding of that data is a new and important responsibility for security departments.

Internal and External Big Data

The pace of business today — and the increasing degree of change to the whole set of business risk factors — warrant very close attention to a great many internal and external factors affecting business. Data are the foundation of decisions for an organization and its activities, and both internal and external data are important in business decision-making and planning.

Mike Page, vice president of client services and technology for Blueocean Market Intelligence, reports: "The retail industry has been particularly successful when it comes to integrating internal and external data sources, especially when it comes to geography, time and product. If a business has any of these components, then a wealth of free big data awaits. For example, a large restaurant chain wanted to incorporate store-level data with every type of survey research available to build a picture of what success looks like."[14] Similarly, security-relevant external data is also available.

The Growing Availability of External Security Data

Among the categories of companies providing external data are cybersecurity threat intelligence providers, who utilize *big data cybersecurity analytics*. Aamir Lakhani, a leading senior security strategist, describes big data cybersecurity analytics as follows[15]:

> ***Big Data Cyber Security Analytics*** *is the collection of data sets that are large and complex. The size, diversity and complexity of this data make it difficult to process using traditional applications. Traditional security tools are great at processing similar data sets with easy-to-understand relationships.* ***Big Data Cyber Security Analytics*** *specializes in organizing solution sets that are unstructured and that are more difficult to find and correlate. Examples of this are finding how Twitter hashtags, road construction or weather affects retail sales of a specific product. Building and analyzing these types of relationships with traditional databases is very difficult, if not impossible.*
>
> *The challenge with Cyber Analytics is how to capture, process, store, search, share, analyze and present large and diverse amounts of data in a visually compelling format.*

Currently there are many more sources of cybersecurity threat intelligence data than there are sources of physical security threat intelligence data. As big data technologies continue to develop, as the capabilities for extracting threat intelligence from big data sources continue to improve, and as more external big data sources emerge, the availability of threat intelligence data for physical security will increase, and the amount of actionable intelligence will increase.

In the near future organizations will be able to subscribe to cloud-based services, which, for example, will identify the safest and most reliable modes of public and private transportation

14. Page, Mike (7 Oct 2015); "A Holistic Approach to Combining Internal and External Data", data-informed.com, 18 Aug 2016, <http://data-informed.com/a-holistic-approach-to-combining-internal-and-external-data/>

15. Lakhani, Aamir (24 Feb 2016); "Understanding and Preventing Cyber Fraud And Cyber Attacks With Advanced Big Data Cyber Security Analytics," doctorchaos.com, 22 Aug 2016, <http://www.doctorchaos.com/understanding-and-preventing-cyber-fraud-and-cyber-attacks-with-advanced-big-data-cyber-security-analytics/>

to evacuate personnel from at-risk overseas locations back to home or office, book the travel upon approval, and monitor the progress of personnel travel in real time. On a smaller scale, such services will be able to monitor local facility conditions and if large protests erupt at a nearby event venue, recommend alternate routes home for facility personnel and even send the routes to the GPS applications on their cell phones. Today, implementing such capabilities would be cost-prohibitive, but in the coming years such services will be both commonplace and affordable.

Physical Security Systems and IT

It is important to remember that at the same time as corporate security departments are being impacted by technology changes, the same basic technology changes have already been affecting corporate IT departments. The implications of this include:

- The information technologies underlying security system improvements are known and understood by IT departments, and are being funded for business applications. The validity and value of these technologies are already understood by organizations.
- Also, the more that information technology is utilized by physical security products, the more security industry vendors can focus on application-specific training and education, leaving IT industry training to cover the underlying technologies. This is the general direction in which security technology is headed.
- The more "mainstream IT" security systems become, the easier collaborations with IT departments will be.
- Reliability and fault tolerance for the information technologies that underlie security systems are improving even as costs keep dropping. This means that security technology will continue to become both more reliable and more affordable, and security systems will be more conformant to IT requirements for the continuity of critical computer and information systems.

Technology Impacts on IT Departments

The primary disadvantage from a security department's point of view is that many corporate IT departments are being heavily impacted by technology's rapid advancement, and the related burdens on IT personnel reduce their availability for support of security technologies. When asked, at the 2014 Gartner Data Center Conference of December 2014, about the effects of digital disruption on General Electric, Chris Drumgoole, chief operating officer of GE's cloud division, said, "There is really not a single thing that we do in IT, today, that we'll do the same way two years from now. I struggle to name a single process within our organization that isn't going to change dramatically over the next two years or three years."[16]

Additionally, as the utilization of information technology in security industry products and services becomes more mature and more conformant to IT standards and practices, collaboration between security departments and IT departments will become easier and more

16. Gartner, Video (December 2014): "Rethink Your IT and Cloud Strategy, Transform Your Business", gartner.com, < https://gartner.mediasite.com/Mediasite/Play/b43aac637af147799bf7ea7a2ea734401d>

productive, and IT evaluation and approval of security system devices and applications will be simpler and easier. Security department technology deployments will be better aligned with IT department standards, practices, guidance and governance. For example, the handling of security systems data will be consistent with IT policies, and the data security of security systems will be easily auditable as part of IT's normal practices.

Aligning Security's Big Data and Privacy Practices with IT

As mentioned earlier, the deployment, operation and maintenance of information-technology-based security systems and devices should be done in accordance with an organization's IT policies and practices for critical systems, and in compliance with regulatory and organizational policies regarding the privacy of personal information and the confidentiality of company information. The following aspects of security systems deployment, including big data usage, are relevant to IT policies and procedures:

- Systems network usage and networking requirements
- Information security, including cybersecurity and privacy safeguards
- Network and systems logical access control
- Product and system acceptability for deployment across the corporate network (IT's approved product list)
- Data governance for data generated by security systems
- Data backup for all system databases, device configurations, analytics systems configurations, and analytics machine learning and metadata databases

Security departments can prepare themselves to benefit from big data by learning from enterprise big data initiatives and understanding the key roles in them. This will provide security practitioners with insight into big data project elements, and give them a vocabulary with which to discuss project topics and accurately report on project progress and status.

Part Two: Privacy

Security practitioners have long been experienced in dealing with private matters confidentially. However, video analytics, big data technologies, and soon IoT will expand the number and type of privacy issues that must be dealt with.

Personally Identifiable Information

All organizations have personally identifiable information (PII) in their human resources and legal departments, and sometimes in the health department if the organization has one. Regulations as well as company policies and practices determine how such information is to be handled.

PII is not new to the security function, as investigations, access control records and video recordings all contain PII. Now, advanced video analytics metadata will also include PII. Fortunately, by establishing good data stewardship for the video analytics data shared with other business functions, the organization's data governance program can oversee data stewardship of the other business functions, relieving the security department of the

responsibility for data that is delivered into other hands. However, security departments must take the initiative to see that good data stewardship is established and that the organization's data governance program properly incorporates the shared data into its governance oversight.

Privacy and Big Data Risks

Surveillance video analytics is a big data application providing increasingly greater value as video analytics capabilities continue to advance, and as more and more data that is generated falls into categories that require privacy protections.

Privacy Concerns for High Resolution Video and Analytics Metadata

Before the days of video analytics, value was extractable from the video images only by human observation and analysis. Searching video data required a human individual to closely watch and analyze the video while playing it back. Today, video analytics software extracts information out of each video image, identifying the various objects in the image and their characteristics, such as color, shape, and size, plus other information such as type of object (i.e. a person or a vehicle) and direction and speed of movement. A generally accepted privacy standard has been that whatever a person could observe about others with the unaided eye while going about his or her normal activities would not have privacy expectations. That thinking could easily be applied to security video in the days before security video was recorded, and when live human observation was the only means of video evaluation. PTZ cameras changed that situation, especially since zoomed-in video could be recorded. There are some state statutes (often referred to as "voyeurism" or "peeping Tom" laws) that prohibit the use of telescopic lenses to look inside residential dwellings, including telephoto lenses on still photo or video cameras.

Thinking along this line has evolved even further, because even without telephoto lenses, high megapixel video images contain more information than is discernible by the human eye, raising new privacy concerns.

This is further compounded by big data and the fact that data analytics can evaluate quantities of data at rates that exceed what a human is capable of. Thus, organizations come into possession of personal information that in previous times would be considered private, because a human being could not discern it from direct observation or by personal access to typically available public databases.

These are reasons why access to video systems and video data needs to be closely controlled by policy-based means, with audit trails that are easily auditable and actually do get audited. In the event of an unauthorized release of data considered to be private, such controls help make the assertion that the organization was not negligent in its handling of video recordings and related video analytics data.

Security Watchfulness: A Corporate Duty

In any company, management is charged by ownership with the care of company assets. Thus, management is responsible for properly utilizing assets, as well as addressing risks to company assets. Management delegates to security the management of security risks, including physical

risks to company assets. This is the root and source of the obligation of security to utilize appropriate people, processes and technology measures for asset protection.

Video technology cost-effectively enables security "watchfulness" with regard to critical assets (people, material and critical processes) to help protect them. Additionally, the facility security responsibility includes maintaining a safe and secure workplace. Video surveillance helps security keep an eye on personnel safety issues, including conformance to safety policies and procedures. Security and safety measures help protect facility personnel from risky or harmful situations. That includes protection from the occasional bad behavior of other employees or visitors. Thus, from a safety and security perspective, video plays an important role in helping to protect and safeguard facility personnel.

Advanced video analytics can now detect aggressive behavior and other suspicious or dangerous activity that may not otherwise be knowable in time for preemptive response by security personnel. Thus, such video analytics support the duty of watchfulness with regard to personnel and property protection.

While there are many valid reasons for utilizing surveillance video, there is a balance to be maintained between the comfort level of employees and visitors (including the corporate culture element of being respectful to people), and the need and obligation to protect critical assets and the critical processes that depend upon them. People have varying considerations about privacy, and about the potential for misuse of surveillance video. Even when people are aware of and have agreed to video surveillance, and thus privacy expectations do not apply, such agreement is given with the expectation that the recorded video data will be handled responsibly. The agreement to be subject to video recording for facility security purposes does not provide agreement for such video to be posted on YouTube, Vimeo or elsewhere, or to be distributed to other facility personnel.

Even when privacy issues don't apply, failing to securely restrict access to security video can result in video misuse and damage to the trust and morale of personnel and to company reputation.

Business, Legal and Regulatory Considerations

There are a number of key considerations with regard to the use of video that relate to its value to the business and to personnel, and to the civil rights and privacy expectations of personnel. *It is important to ensure and to document that current and planned future use of security video technology does not conflict with individual civil rights, federal or state regulations, or company business policies.*

Regulatory Issues

Federal civil rights and state privacy regulations prohibit the use of video technology where these violate the privacy rights of individuals. Where state laws provide stronger constraints than federal civil rights and rulings do, the state constraints apply. Currently, about a dozen states have laws regarding covert surveillance video that address the issue of individual privacy rights.

Video Data Protection

Due to the value of security video data, it is important that appropriate data protection measures are established. Such measures should be consistent with those typically applied by the organization to critical business electronic data systems to assure the confidentiality, integrity and availability of the data. To do less opens an organization up to accusations of negligence if problems arise due to unauthorized access to video systems and data. Video retention is also an important issue. Statutes allow slip-and-fall claims to be filed up to two years after the event, which makes it paramount to establish a clear video retention policy that takes this and other statutes and court rulings into account when determining what retention requirement will be applied to various cameras.

Individual Privacy Expectations

A "reasonable expectation of privacy" is the central issue with regard to individual privacy rights. Video should not ever be used in facility areas where an expectation of privacy would exist, such as changing areas, locker rooms, shower areas or restrooms. Courts have ruled that work areas such as lunch or break rooms, and any area which is open to occupancy or viewing generally by facility personnel, do not have a reasonable expectation of privacy. In particular, legal precedent states that "when an individual enters into an employment situation with high security requirements, it becomes less reasonable for her to assume that her conduct on the job will be treated as private."[17]

Individual Legal Rights

Specific individual legal rights apply to situations where *covert video surveillance* is used in areas where a legitimate expectation of privacy exists. Suspicion of illegal behavior in the workplace does not nullify an individual right to privacy. Where covert video is used to investigate suspected criminal behavior in areas having legitimate privacy expectations, probable cause and search warrant requirements apply, even if the surveillance has been requested or is being performed by law enforcement personnel. The following case description[18] provides an example of how complicated the considerations can be. (Underline emphasis has been added for this paper.)

> *A covert surveillance case from the 9th Circuit Court of Appeals establishes the important distinction between administrative and criminal investigations. In [the] Taketa [case], a DEA agent reported to her supervisor that another agent, Taketa, had shown her how to turn an authorized pen register into an unauthorized wiretap to record conversations. Agent Taketa shared an office at McCarran Airport in Las Vegas with Thomas O'Brien of the Nevada Bureau of Investigation. O'Brien was also involved in the illegal wiretapping. The DEA began an internal investigation and entered the airport office using a key from headquarters. The investigators examined the feasibility of putting in a covert camera the next time the Taketa sought authorization for a pen-*

17. Ryan, Jack (2008); "Covert Video Surveillance", patc.com, 16 Aug 2016, <http://www.patc.com/weeklyarticles/covert-video-surveillance.shtml>

18. Ibid.

register. In its review of this first entry, the court applied the standards announced in O'Connor v. Ortega in holding that the first entry was reasonable.

In May, Taketa sought such an authorization and the investigators returned to the airport office. Unable to find the recording equipment, the investigators had to force open O'Brien's door with a plastic card to gain entry. The investigators found the recording device. They then placed a covert camera in the office. This investigation led to the arrests of Taketa and O'Brien. In its review of the covert surveillance, the court held that the video evidence had to be suppressed. In so holding, the court determined that since O'Brien's office had been locked, investigators should have obtained a warrant before entering. In addition the court concluded that once the investigation changed from an internal investigation to a criminal investigation, the standards from O'Connor v. Ortega no longer applied, rather the more stringent standards of probable cause and a search warrant were required.

Use of Covert Video Surveillance

Covert video surveillance is the unannounced use of hidden video surveillance. Court rulings have established legal precedents regarding the acceptable purpose, scope and nature of covert video surveillance with respect to personal privacy issues. An organization's security policies and procedures should document the following aspects and requirements for the use of covert surveillance:

- Allowable uses
- Managerial approvals required, including:
  - a decision-making task force or committee of non-biased personnel to mitigate possible claims of prejudice or unfairness that can arise from having a single-decision maker with regard to the use of covert video surveillance
  - a specific time limit for the use of covert video surveillance, as well as criteria by which the surveillance can be ended early due to the surveillance having achieved conclusive evidenceInternal reporting requirements
- Legal consultation requirements
- Protections that must be established for:
  - the secrecy of the video data
  - the privacy interests of any affected personnel
  - requirements for consultation with legal counsel
- Special considerations when the nature of an investigation is criminal in scope or becomes criminal in scope
- Sharing video data with law enforcement

These documents should be reviewed by an attorney initially and whenever revisions are made that reflect the use or application of covert video surveillance.

Control of Access to Video

Written security policy should exist regarding the sharing of live or recorded video with authorized non-security personnel. These should reference relevant aspects of an organization's

codes of conduct, policies regarding the use of electronic media and policies regarding the protection of company information, as often these can establish a basis for the video-specific policies.

Potential Misuse of Video

In most organizations, misuse of security video is an unlikely scenario due to the limited number of personnel who can access video, and the nature of their positions in the company. However, organizations that don't have written policies and procedures on the acceptable use of video should establish them.

In particular, is it important to protect the company against liability exposure from potential employee misuse of live and recorded video. This can be done by establishing appropriate company and departmental policy that restricts the use of video to authorized purposes, and requires signed statements (from employees authorized to access video) that the policies have been read and will be adhered to. Court decisions have held that organizations must periodically provide refresher education and review of such documents, which most organizations preform annually.

Verifying Privacy Protection

Establishing regular periodic audits of the privacy protections in place for surveillance video is important to help safeguard the organization against claims of insufficient data security and negligence, if an unauthorized release of privacy-restricted video information does occur.

Government Use of Video in Public Spaces

The government use of surveillance video in public is subject to an additional set of privacy and legal considerations. These also apply to the many state initiatives for public-private partnership around the use of personal and company outdoor security surveillance video. For example, on November 9, 2015 the State of New Jersey enacted legislation[19] authorizing the establishment of law enforcement agency registries of citizen-owned and company-owned outdoor video surveillance cameras. The law allows New Jersey municipalities to establish *voluntary programs* for registration of outdoor video cameras, to establish a database for law enforcement use to be able to quickly determine what sources of security video recordings might be available to them, without their having to canvass the neighborhood and go through a process of obtaining the required permissions to obtain the video recordings of interest.

The private outdoor video surveillance camera registry requires:

- The name of the camera owner
- The owner's address and phone number
- The street address where the camera is installed
- The number of cameras

19. State of New Jersey (2015); "New Jersey General Assembly passes camera registration bill," securityinfowatch.com, (5 Feb 2015), <http://www.securityinfowatch.com/news/12041870/new-jersey-general-assembly-passes-camera-registration-bill>

- Identification of the outdoor areas recorded by the camera
- How the footage is stored or saved

As of today, several dozen city and county police departments in New Jersey have implemented such registries.

Policy Considerations for the Use of Video in Public Safety

The legal and privacy issues involved with government use of outdoor video surveillance in public areas have been very thoroughly addressed by a 76-page guidance document, *Policy Considerations for the Use of Video in Public Safety,* published in June 2016 by the U.S. Department of Homeland Security Science and Technology Directorate[20].

Any private sector organization can look to this document for guidance in preparing for and documenting the use of surveillance video for outdoor event venues. In particular, the preparation of a Privacy Impact Statement should be considered, as it will establish a basis for answering questions about outdoor video surveillance usage should this ever become a public relations consideration, or if an incident occurring during an event becomes newsworthy.

The following section contains an outline of the recommended contents for a video surveillance Privacy Impact Statement, taken from page 16 of the document.

Privacy Impact Statement Contents

Whether or not a privacy impact statement is required, users seeking to implement a video system would benefit by considering the following essential elements in the design and operation of a video system. Any written policy should explicitly discuss:

- Why video is being collected and retained;
- Whether cameras will be covert (hidden) or overt;
- Whether there will be notice given to those in the area (i.e., with signs);
- How the images will be used;
- What analytics (i.e., automated systematic computational analysis), if any, will be applied to the video data;
- Whether attempts to identify individuals in the video data will be made systematically or on a case-by-case basis;
- What other information will be combined with the video as part of processing;
- Who is authorized to view the images and the processed data;
- How long the video will be retained under normal circumstances;
- What measures will be necessary to block or override automated deletion;
- Whether the results of analytics are stored directly with the video or stored elsewhere;
- Whether additional privileges are required to access the results of video analytics;

20. Department of Homeland Security (24 June 2016); "Policy Considerations for the Use of Video in Public Safety," dhs.gov, (18 Aug 2016), <https://www.dhs.gov/publication/vqips-policy-considerations-use-video-public-safety>

- What procedures will be followed in order to disclose the videos to others, both inside and outside the organization; and
- What procedures will be followed prior to public disclosure of video data.

The Importance of Documentation and Verification

Data governance plays an important role with regard to privacy issues relating to security information, especially for video image and video analytics data. For most medium and large organizations, there are existing policies, processes and procedures in place that govern planning and documenting the creation and usage of the information systems, establishing data stewards for their data, and regularly auditing to verify that the deployment of the technologies and their usage has been established and maintained as documented. Where these do not exist, sensible approaches should be adopted to achieve the results just described.

Conclusion

Although much of the information in this paper may be new to many readers, over time it will become common knowledge and common practice for most organizations. Regarding big data and its related privacy considerations, the security industry and its customers will be walking down paths already walked by many IT departments and many users of business information systems. It is heartening to realize that at this point in time, the promises offered by big data, and the benefits that many organizations have already been receiving from it, are truly becoming available to organizations of any size and those responsible for their security.

Ray Bernard (RayBernard@go-rbcs.com) is president and principal consultant of Ray Bernard Consulting Services (www.go-rbcs.com).