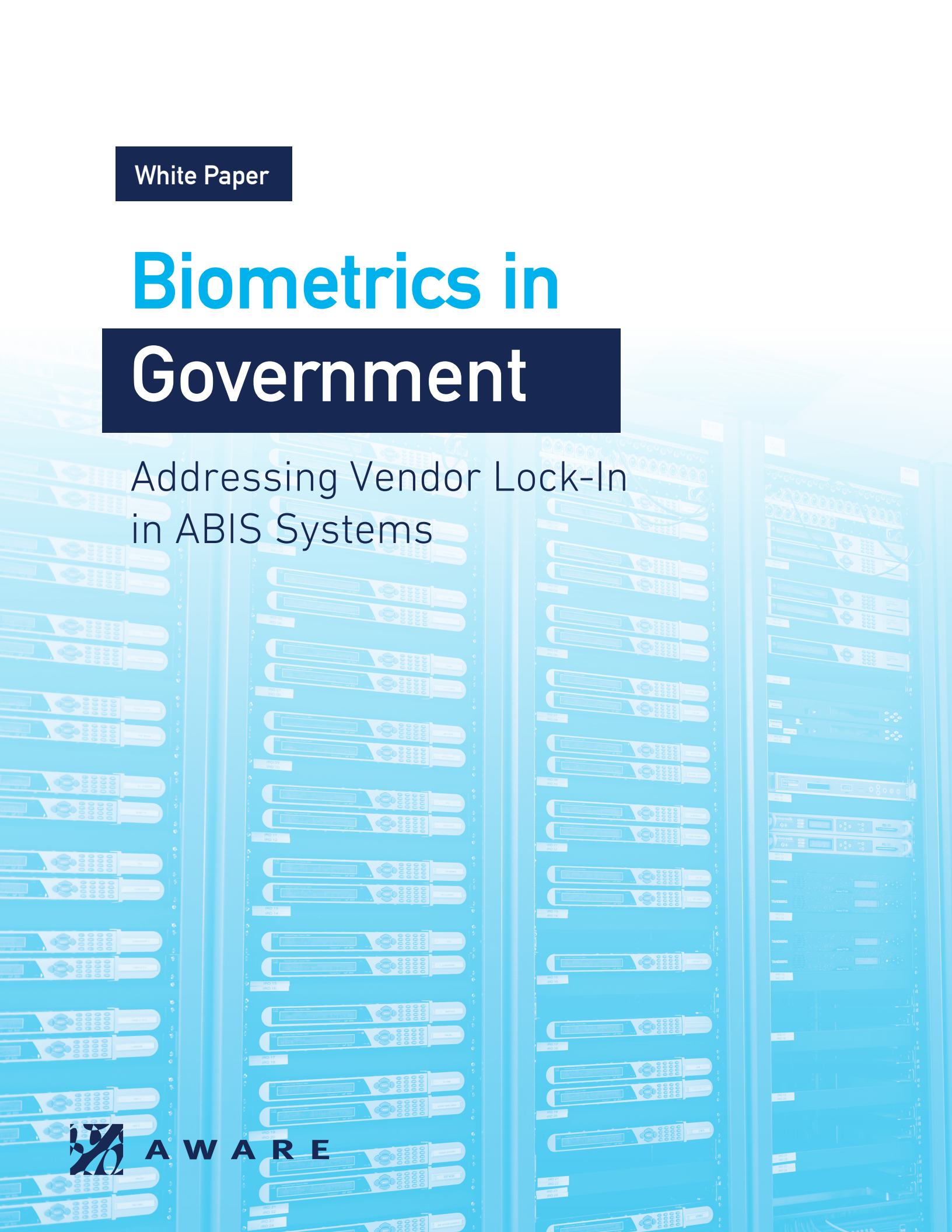
White Paper

Biometrics in Government

Addressing Vendor Lock-In in ABIS Systems



AWARE

Introduction

Preventing lock-in in biometric systems has never been more achievable.

"Vendor lock-in" is a term that describes when a technology vendor imposes switching costs upon their customer—intentionally or otherwise—to make it unattractive for them to replace their installed products. It's done by designing and deploying a system in such a way that makes it exceedingly difficult, risky, or expensive to replace part or all of the system. The effect is that the vendor can earn a virtual monopoly within that account on future product and service revenue. The phenomenon is universal, but vendor lock-in has occurred in the biometrics realm, typically involving deployment of an AFIS/ABIS (automated fingerprint/biometric identification system).

What is vendor lock-in?

There are many ways that a product can be designed or installed that increase the likelihood of vendor lock-in. A system prone to vendor lock-in tends to be "monolithic" in design, with proprietary protocols to communicate between subsystems, or lacking subsystems altogether. They are difficult to integrate with other systems and difficult to replace or upgrade in an incremental fashion. In basic terms, vendor lock-in arises when a system is closed.

Data communication standards and the compliant implementations that adhere to them are intended to make systems open. They are comprised of rules that dictate how two systems or subsystems must interoperate. These rules fall into two categories: 1) connectivity protocols and 2) data interchange formats. The first establishes a means for the systems to request, send, and receive data. The second defines what the shared data is, where it is, and how to interpret it.

While connectivity is relatively simple to establish, formatting and sharing data that can be properly interpreted tends to be more complicated. This is where vendor lock-in can become a particularly challenging problem. If data is not defined and formatted in a way that's common to two communicating systems, they cannot effectively interoperate.

Vendor lock-in: an analogy

A helpful analogy found outside the digital realm is human communication. For two people to meet and converse (unaided by technology), they both must make arrangements such that they can each speak to and hear the other party in real time. They need to meet at a time and location that is convenient for both. If a meeting isn't possible, they might choose to correspond by mail, which would require exchange of addresses. Arranging such communication is not terribly complicated but requires some interaction and planning. Let this represent our connectivity protocol.

Now imagine that these two people do not speak the same language, or even use the same alphabet. In this case, the two people will need, at a minimum, to use some sort of translation dictionary. That dictionary will only have to include those words needed in their discussion. This is our data interchange format. But does such a dictionary exist? Where to get one? What words will it include? It is a difficult way to have a conversation, but without a dictionary, it is virtually impossible.

Computers, however, are quite good at using dictionaries…

If data is not defined and formatted in a way that's common to two communicating systems, they cannot effectively interoperate.

Computing systems – proprietary vs. standards-based

Disparate computing systems similarly need to establish connectivity to communicate, and data formats to structure and interpret data. Technical standards written collaboratively are intended to achieve this.

Different components of a system, such as a client application and a server application, or a client application and a peripheral device, need to communicate with one another. They do so when they are both designed to establish connectivity in the form of data requests and responses with one another in a common way. In some cases, this communication will be proprietary; i.e. it will use a language all its own. A system that uses proprietary, closed protocols to communicate and to exchange data requests and responses is prone to vendor lock-in. Proprietary communication protocols will require that the client application and the server work only with each other. If a need or desire to upgrade or replace the server technology comes about, all client and server technology must be replaced at the same time. For a large system with hundreds or more clients in use, this makes replacement of the back-end system substantially more costly. This is a generic version of vendor lock-in.

This is in contrast to a modular, standards-based approach that allows different client applications from different vendors to operate with the same back end system and a variety of peripheral devices. Much of the work in standards bodies is around specifying universal connectivity protocols and data interchange formats. The explicit intent of this work is to prevent vendor lock-in and technology monopolies, and to facilitate an open market that drives competition and innovation.

Amazing progress has been achieved in the last decade or so towards preventing vendor lock-in. Technologies such as REST architectures that utilized standardized application programming interfaces (APIs), and web browser standards that make extraordinarily powerful applications possible without any client-side code required at all. But these systems still require a common language to format and interpret data being exchanged between two systems.

Biometric systems - what is an ABIS?

An ABIS is a biometric search platform used by a government agency for law enforcement, border management, or civil ID programs to establish or confirm a person's identity, or to detect an attempt to misrepresent identity. At a minimum, a complete biometric search solution requires:

1) hardware peripherals and software running on client workstations or mobile devices, which are used to collect and/or analyze biometric data, and

2) server software that processes, searches, stores, and exchanges the biometric data.

An ABIS is often integrated with systems of other government entities. A criminal ABIS is used by law enforcement for investigations, and adds the ability to capture, analyze, and search latent fingerprints found at a crime scene or facial images taken from surveillance video.

As with other types of complex computing systems, establishing connectivity between two biometric subsystems is less complex than standardizing the data that gets exchanged so that it can be properly interpreted and processed.

Preventing vendor lock-in in biometric systems in three steps:

1. Preserving an archive of raw biometric images and data

But there's a characteristic particular to biometric systems that makes them even more susceptible to vendor lock-in, which is the fact that the biometric templates that are used by computers to compare biometric data are always inherently proprietary. Every ABIS provider has their own algorithms that extract the features of a raw biometric sample (e.g. a face, fingerprint, or iris image) to create a "template", and then compare those templates as part of a search. These algorithms are an important part of how biometric algorithm providers differentiate their products; they can be optimized for speed, size, or accuracy, for example.

So vendor lock-in can happen in a biometric system if it does not preserve the original raw biometric images, or does not foster unfettered access to them. This is because for the system owner to replace the ABIS, or even just add and fuse a new algorithm, new

templates must be generated for all the previously collected biometric data using the new algorithm. If the raw data isn't available or accessible, the new algorithm can't be used with the old biometric data, which could include millions or even tens-of-millions of biometric records.

So the most obvious way to prevent vendor lock-in in biometric systems is to ensure that the raw biometric data is preserved in such a way that it can be used to generate new templates. Conversely, limited access to raw biometric data is also the most obvious sign that vendor lock-in might be a problem.

But it isn't enough to have access to the original biometric data; it must be of the original image quality in terms of compression ratio and resolution; it needs to be compressed in a standard-compliant way; and it must be stored in a data structure along with accompanying metadata that can be interpreted by the new biometric system.

2. Ensuring standards-compliance of biometric data

Standards have been critical to the successful use of biometric systems for decades. Standards like WSQ and ANSI/NIST-ITL were initially developed by the US law enforcement community, but they have proven essential in biometric systems around the world. The WSQ standard specifies how fingerprint images are to be compressed (and decompressed). The WSQ algorithm is designed specifically for fingerprint images, which due to their nature are difficult to compress efficiently. The ANSI/NIST-ITL standard and its derivatives specify the data format with which biometric images and metadata can be stored and exchanged.

The standards are powerful; two systems designed to adhere to WSQ and ANSI/NIST-ITL standards can exchange biometrics by email if necessary; in fact, SMTP (Simple Mail Transfer Protocol) is a common way for biometrics to be exchanged between disparate parties. Without WSQ and ANSI/NIST-ITL, it's extraordinarily difficult to exchange biometric data.

3. Employing a modular architecture

Implementing a modular architecture that allows subsystems to operate independently and exchange data in standardized formats is of paramount importance. Instead of a monolithic platform with closed, proprietary connections to subsystems, a system built around a standards-based "hub" or "service bus" can ensure that all subcomponents can communicate with all others. A hub can greatly simplify a biometric network by consolidating communication and processing functions. (Consider how airline hubs expand the number of cities we can visit.) With such an architecture, any subcomponent can be replaced or upgraded with its equivalent.

For example, a biometric matching system communicating with client applications or devices through a standards-based hub ensures that those clients and the matching system operate independently, can be procured separately, and can replaced or upgraded independently in the future.

Preventing lock-in in biometric systems has never been more achievable

Modular, open architectures are the norm today, driven in part by remarkable advances in services-oriented, cloud-based computing and browser technologies. They have transformed stove-piped enterprise computing platforms into distributed computing networks, software as a service, and microservices. The enterprise service bus is a product of this evolution, and modern biometrics systems are using them.

But all this cloud technology does not alone solve the vendor lock-in risk. Establishing connectivity between systems is just the first step. The complexity is in formatting and interpreting the data. What data is where? What does it mean? Biometric systems are no different. Those that that use ANSI/NIST-ITL standards-based data interchange formats and certified versions of WSQ are far more likely to be protected from vendor lock-in risks.

Please contact Aware for more information about how we can help you migrate your ABIS and biometric data to a solution that is open, flexible, extensible, and future-proof.

www.aware.com/contact

AWARE