Dataversity – Preventing Nation-Wide Security Breaches with Semantic Standards

Last year's Equifax security breach provides a valuable lesson for organizations seeking to protect digital assets and is worth examining closer for three reasons. With intruders accessing personally identifiable information such as social security and driver's license numbers for over 100 million Americans across the country, it's one of the most severe breaches of its kind. It's the company's second major security penetration of the year and the third in the last two years, attesting to the frequency of breaches today.

But, more significantly, this breach suggests how to mitigate the effects of such occurrences. A New York Times article citing Gartner fraud analyst Avivah Litan states: "Equifax should have multiple layers of controls" so if hackers manage to break in they can at least be stopped before they do too much damage."

Supporting external security measures with strict internal ones can prevent intruders from accessing valuable information. A Semantic Standards-based approach characterized by role-based access to data, key value triple attributes, and unerring Data Lineage provides multiple controls to check intruders. The combination can greatly reduce security breach damage, if not preclude it altogether.

Role-Based Access

The implementation of role-based data access is perhaps the most distinct way standards-based technology underpins security. Semantic statements (called triples) can denote who can access which types of data and how. Role-based access is a fundamental way of ensuring proper data governance measures

are followed so only predetermined users are privy to certain data. It’s an internal security filtering mechanism for narrowing the scope of users able to view specific datasets.

Typically, user-based access is determined by one’s position or responsibility. Still, the factors by which triples can provision data access are virtually limitless, enabling organizations to write them in accordance to security protocols most relevant to a particular dataset. This methodology provides a crucial initial layer of internal security, supplementing whichever external measures are employed.

Additional Attributes

Increasing internal security with triples is enhanced by bestowing key-value attributes to these semantic statements. The benefits to augmenting user-based access with these additional attributes are manifold. Firstly, this layer of security occurs deep within the database (or triple store) in which the data reside. Thus, it’s a granular security layer with role-based access and external security measures on top, compounding the overall complexity required to access data. Secondly, organizations can input as many key-value attributes as needed to fortify data identified as triples.

With this approach, all data is converted into triples and linked together in a Semantic Graph supported by standardized models and classifications. Even if an intruder is able to penetrate external security and bypass the role-based access layer, he or she must still have the requisite key-value pairs to reach the protected data. Moreover, the attributes for the key values can be even more arbitrary than the triples for security filtering – which is another way they strengthen an organization’s overall security model. Users dictate whether certain key value attributes enable permission according to a specified range or an exact numerical value.

For example, the latter might involve a security clearance level above five, whereas the former functions as a revolving pass code. A combination of these types delivers robust protection. Organizations adopting this approach have a fine-grained, flexible security model in which users must fulfill all of the requirements – the key-value pairs and the security filters – to access data via queries. Lastly, this paradigm adds the JSON standard to the RDF standard, which is a novelty in itself with which unauthorized users must contend.

The Data Trail

An added security measure which proves useful in the event of breach attempts is the data lineage capabilities of the underlying Semantic Standards approach. Because data are linked together on an RDF graph and connected with evolving models, users can easily determine the various facets of a specific dataset's lineage. Organizations can ascertain any changes that data underwent (such as transformation or those stemming from application usage), as well as which users manipulated data.

These provenance capabilities can also yield insight into who viewed data, whether they were replicated, or even how they were moved. In the event of a breach, this fundamental understanding of data's traceability informs users of how security was penetrated and what activities occurred as a result. Users can readily determine which data were compromised, possibly gaining insight into important characteristics of those perpetrating the attack.

Internal Security Measures

Unfortunately, security breaches are becoming more commonplace. Recent breaches of the SECverify this fact alongside Equifax's troubles. Organizations must protect data with security characterized by increasing layers of difficulty. Far too many contemporary events prove that

external security is just the beginning. The greater degree of fortification takes place within the enterprise via role-based access, triple attributes, and Data Lineage.