www.jusletter-it.eu

Ann Cavoukian / Michelle Chibba / Graham Williams / Andrew Ferguson

The Importance of ABAC to Big Data: Privacy and Context

There is little doubt that Big Data is an increasingly important topic as more and more of our data is digitized. Large organizations were quick to realize the enormous potential of correlating vast amounts of data. As with many technology advances, there are issues to be addressed — including the significant potential for Big Data analysis to flaunt privacy rules and expose the analysts to liabilities. This paper considers the privacy issues relating to Big Data analysis and investigates how emerging attribute-based access control technology can assist in protecting against unauthorized access to personal data in a Big Data context.

Category: Articles
Field of law: Big Data, Open Data & Open Government; Data Protection
Region: Kanada

Citation: Ann Cavoukian / Michelle Chibba / Graham Williams / Andrew Ferguson, The Importance of ABAC to Big Data: Privacy and Context, in: Jusletter IT 21 May 2015

Contents

1 Introduction

[Rz 1] With the emergence of Big Data, the Internet of Things, Open Government and the evolving technologies enabling these trends, organizations are now extracting value from the massive amounts of digitized data available to them. Personal information is invariably part of this mix. We believe that this trend is based on amassing and linking large unstructured and structured datasets, as well as the fact that greater sharing of data ignites a renewed interest in the applicability and evolution of access controls. This paper finds its focus knowing that integral components of an organization's privacy and security program are appropriate safeguards to ensure that personal and confidential information in its custody and control are not compromised through inappropriate access, and that regulations within the jurisdictions in which they operate are not breached. The goal from both a security and privacy standpoint is to ensure that authorization to access information is on a legitimate «need to know»[1] [2] basis.

[Rz 2] The purpose of this paper is two-fold: First, to provide an overview of the significance that access controls play in data protection and data sharing, particularly in a Big Data and multi-stakeholder data sharing environment, and acknowledge the more commonly used role-based access approach. Second, we would like to introduce attribute-based access control technology as a Privacy by Design solution for preserving access to personal data and provide a high level roadmap of essential factors that organizations should consider prior to implementation.

[1] Glossary Of Key Information Security Terms. 2011. Washington, DC: National Institute of Standards and Technology (NIST). http://csrc.nist.gov/publications/nistir/ir7298-rev1/nistir-7298-revision1.pdf.

[2] See also guidance provided by Canada's Federal Privacy Commissioner's Office. «Employees» access to personal information should be limited to what they need to know, particularly when this information is sensitive.» https://www.priv.gc.ca/resource/fs-fi/02_05_d_60_tips_e.asp.

2 Designing in Privacy and Security by Default: A proactive, POSITIVE-SUM Approach

[Rz 3] The concept of Privacy by Design (PbD) was developed over a decade ago, and over the years, has achieved global acceptance.[3] It emphasizes respect for user privacy and the essential requirement to proactively build-in privacy as a default condition to information technologies, business practices and networked infrastructure. Not only does this framework for privacy in the 21st century address privacy imperatives, it also commits to preserving doubly-enabling functionality and business objectives — otherwise known as «positive-sum» versus the win/lose approach that balances competing interests by taking a zero-sum perspective. Moreover, with the growth and complexity of information communication technologies (ICTs), a reactive regulatory compliance model to safeguarding privacy is clearly insufficient as demonstrated by some recent high-profile data breaches that have resulted in significant remedies being sought by both staff and members of the public whose privacy was violated. Clearly a more comprehensive «by design» approach is warranted.

[Rz 4] The 7 Foundational Principles of Privacy by Design are as follows[4]:

1. Proactive not Reactive; Preventative not RemedialThe Privacy by Design (PbD) approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy invasive events before they happen.
2. Privacy as the DefaultPrivacy by Design seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. No action is required on the part of the individual to protect their privacy — it is built into the system, by default.
3. Privacy Embedded into DesignPrivacy by Design is embedded into the design and architecture of IT systems and business practices. The result is that privacy becomes an essential component of the core functionality being delivered.
4. Full Functionality — Positive-Sum, not Zero-SumPrivacy by Design seeks to accommodate all legitimate interests and objectives in a positive-sum «win-win» manner, avoiding the pretense of false dichotomies, such as privacy vs. security, demonstrating that it is possible to have both.
5. End-to-End Lifecycle ProtectionPrivacy by Design, extends data security throughout the entire lifecycle of the data involved, from start to final destruction..
6. Visibility and TransparencyPrivacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification.
7. Respect for User PrivacyPrivacy by Design requires organizations to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options.

[Rz 5] As external and internal threat levels rise, organizations must continue to evolve strategies to protect their data assets. Simply building a defensive «perimeter» around a resource will no

[3] Refer to the International Data Protection Commissioners» Resolution on Privacy by Design, October 2012 [http://bit.ly/1GSRvu4]; the U.S. Federal Trade Commission identification of Privacy by Design as one of three pillars for organizations in addressing consumer privacy [http://1.usa.gov/1DygKEC]; the EU draft privacy regulation that refers to taking a Privacy by Design approach to technology development [http://bit.ly/1wr1pAd].

[4] Cavoukian, Ann. 2011. *The 7 Foundational Principles. Implementation And Mapping of Fair Information Practices*. Toronto: Information and Privacy Commissioner's Office, Ontario. [http://bit.ly/1CqlWJa].

longer be sufficient, thus requiring an offensive data security approach that addresses security concerns as the default condition. Strong security is essential to achieving strong privacy.

[Rz 6] One aspect of security by default requires that access to information, systems and applications be limited only to the data and functionality that are needed for a particular task. «Need-to-Know» is a method of isolating information resources based on a user's need to have access to a particular resource in order to perform one's job but no more. The terms «need-to-know» and «least privilege» express the same idea. Need-to-know is generally applied to people, while least privilege is generally applied to processes.

3 The Evolving Data Landscape: Challenges to Privacy and Security

[Rz 7] Over the last decade, governments have embarked on initiatives intended to foster innovation in technology development, create knowledge-based economies, foster job creation and enhance greater public service accountability and transparency through enhanced citizen engagement in the democratic process. The digital revolution has provided the opportunity to extract value out of tremendous amounts of data that are now available to organizations. Exploiting these opportunities invariably involves data-sharing and raises the issue of regulatory constraints on divulging personal information. In a business context, the trend has been to more open networked enterprises[5] that span the globe thus making it increasingly more difficult to adequately protect critical data assets. At the same time, personal information is taking on an unprecedented online value. Transactional data collected by businesses and cross-industry entities such as media analysts are now seen as a rich source of information for marketing activities. These trends introduce unprecedented challenges to organizations in the areas of privacy and security. As this evolution gains momentum it is incumbent upon organizations to acknowledge that this movement towards greater collaboration, openness and predictability brings with it the imperative of robust data protection and a constant recognition of the privacy rights of individual citizens whose personal data is involved.

[Rz 8] For example, in 2002, the UK Cabinet Office issued a blueprint report on Privacy and Data Sharing: The way forward for public services[6] noting that embarking on such a path requires that «the public should have confidence that public services are taking steps necessary to control access to their information and keep track of who has accessed it. Access controls are one tool for protecting personal data from misuse and audit trails are important tools for both citizens and for public sector bodies in managing information.»

[Rz 9] Consider also the privacy and security concerns associated with Big Data.[7] As technological advances improve our ability to exploit Big Data, potential privacy concerns could stir a regulatory backlash that would dampen the data economy and stifle innovation. Navigating the massive

[5] Cavoukian, Ann. Privacy and the Open Networked Enterprise. Information and Privacy Commissioner's Office. December 2006.

[6] Performance and Innovation Unit. 2002. Privacy And Data Sharing: The Way Forward For Public Services. UK Cabinet Office.

[7] For further discussion on impact of Big Data on privacy, security and identity see for example, Cavoukian, Ann / Jonas, Jeff. Privacy by Design and Big Data. June 2012. Tene, Omer / Polonetsky, Jules. 2013. «Big Data For All: Privacy And User Control In The Age Of Analytics». Northwestern Journal Of Technology And Intellectual Property Vol 11 (Issue 5). http://scholarlycommons.law.northwestern.edu/njtip/vol11/iss. Richards, Neil M. / King, Jonathan H.. 2014. «Big Data Ethics». Wakeforest Law Review. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2384174.

volume of information represented by Big Data requires a new and potentially innovative approach to data collection, storage, sharing and destruction. These efforts must include provisions for privacy protection. By way of example, algorithms can now automatically infer that different digital transactions in different systems are in fact related to the activity of a single person or household. A bank that wants to better serve its customers will be eager to know if a specific customer has three relationships with the bank and has an enormous Twitter following. In the past, identifying the difference between six people, each with one fact, versus one person with six facts was expensive and difficult — something only larger organizations could accomplish. Today, the advanced analytics needed to reconcile related entities over diverse data sets (commonly called Entity Resolution) on a Big Data scale are becoming available to organizations of all sizes. As more data, from more sources, assembles around a single individual, attempts to reliably protect individual identity are more easily compromised.[8]

[Rz 10] The cost of a breach to an organization is not trivial. Research undertaken by the Ponemon Research Institute[9] over the last decade shows that the cost to an organization is in the million dollar plus range, with the average number of records breached at just over 23'000. Extrapolate this to a Big Data environment and the consequences will be far more devastating because of the sheer number of individuals likely to be affected by a Big Data breach.

[Rz 11] Big data solutions often rely on traditional firewalls or implementations at the application layer to restrict access to the information[10] but this approach can hampers access management because of an «all or nothing» limitation; if you can gain access to a protected subnet, then you can access all the information that resides there. Similarly, course-grained access mechanisms tend to encourage placing greater amounts of restrictions on data due to heightened security and regulatory concerns. This hampers the business process that would benefit from the appropriate sharing of the data.

[Rz 12] Given this evolving data landscape, it is clear that more sophisticated access controls are an essential component of an organization's privacy and security program. We believe, on the basis of the challenges to privacy and security outlined above, that there is a compelling case for fine-grained access rights such as Attribute Based Access Controls in Big Data scenarios.

[Rz 13] Indeed, an academic Ram Krishnan at the University of Texas at San Antonio recently noted at an NSF Workshop on Big Data that «Access control is one of the most powerful and fundamental ways of risk mitigation in any application. Interestingly, despite security concerns raised by Big Data applications, there is minimal research in this area·[11]

[Rz 14] Another industry expert, Finn Frisch, noted there is a need to support a shift from a «need-to-know» paradigm to one of a «need-to share». Invariably, this shift increases the need for standards, methodologies, and technologies to take a positive-sum approach that factors in both the ability to protect information and to share that information with those that need it most, while maintaining the privacy rights of the individuals involved. This expert also concluded that

[8] Cavoukian, Ann / Jonas, Jeff. 2012. *Privacy By Design In The Age Of Big Data*. Toronto: Information and Privacy Commissioner's Office, Ontario.

[9] See Ponemon Institute / Symantec. 2013. Cost Of Data Breach Study: Global Analysis. Ponemon Institute..

[10] Lafuente, Guillermo. 2014. «Big Data Security - Challenges And Solutions». Blog. https://www.mwrinfosecurity.com/articles/big-data-security---challenges-solutions/.

[11] See Krishnan, Ram. 2014. «Access Control And Privacy Challenges In Big Data». Presentation, NSF Workshop on Big Data.

traditional access management paradigms were inadequate for meeting this need.[12]

4 Logical Access Control: Expressing Privacy and Context

[Rz 15] A set of business processes and technologies exist, known as Identity Access Management (IAM)[13] or more broadly known in computer security as logical access control[14] that govern and regulate who is allowed access to information stored or being processed within IT environments. IAM relates to how best to protect both the information and the information systems from external and internal attacks, throughout the full life cycle of the data (e.g. unauthorized use, disclosure, modification, or destruction). It is also about an organization managing risks and being prepared, in the event of an incident, to respond to audits, regulatory or law enforcement investigations.

[Rz 16] In the past, access control based on roles was seen as the best way to address such security risks, demonstrate compliance and provide administrative efficiencies. As organizations streamlined and improved the efficiency of their identity management, role management has blossomed across many organizations, creating what has been termed as the «role explosion» (where more roles exist than actual individuals in an organization). Today, the use of roles in providing better access control is common practice but most implementations are specific to platforms and applications. There are, however, a variety of role management software applications that have been developed by specialist IAM vendors, with many subsequently acquired to be part of IAM product suites. These specialist applications have focussed on role discovery, analysis, modelling and design, management and reporting, publishing and dissemination to relying applications.

[Rz 17] On the other hand, policy-based access control or dynamic authorisation management is a more current, forward looking approach in authorization systems. It is not an exaggeration to say that **attribute-based access control** (ABAC) holds the promise of a seismic shift in the development of access control technology. Attribute-based systems provide fine-granularity, high flexibility, rich semantics and other beneficial features like partial authentication and natural support for role-based access control. These features are more advantageous as organizations move to greater collaboration and data sharing across and outside the enterprise.

[Rz 18] These inherent features of ABAC also make it more relevant in supporting privacy and security in a Big Data context. The reason for this relates to «context.» Context is a key factor in Big Data, just as context is key to understanding privacy. By way of a Big Data example, when Google Flu's ability to predict the spread of influenza was found to be overstated, the reason cited was missing information from data subjects on why they were Googling flu-related search terms (they were displaying flu symptoms vs. they were trying to ward off getting the flu). The ability to delve into causation as opposed to only correlation and inference in certain Big Data analysis requires greater access to personally identifiable information. In her book «Privacy in

[12] Frisch, Finn. 2014. *The Identity And Access Management (R)Evolution: Federation And Attribute Based Access Control*. Sweden: Axiomatics AB.

[13] For more on this subject, see Kuppinger, Martin. 2015. «Identity Access Management». Blog. Adaptive Policy-Based Access Management (APAM): The Future Of Authentication And Authorization. https://blogs.kuppingercole.com/kuppinger/category/access-management/.

[14] NIST An Introduction to Computer Security: The NIST Handbook, Special Publication 800-12. (Chapter 17 — Logical Access Control).

Context: Technology, Policy, and the Integrity of Social Life,» Helen Nissenbaum, well-known for her examination of the relevance of context to privacy, provides an account of online privacy in terms of expected flows of personal information, modeled with the construct of context-relative informational norms. She proposes that the key parameters of informational norms are actors (subject, sender, recipient), attributes (types of information), and transmission principles (constraints under which information flows). In contributing to the current debate on the evolving online data landscape specific to the issue of notice, consent and privacy policies, she notes that «We must articulate a backdrop of context-specific substantive norms that constrain what information websites can collect, with whom they can share it, and under what conditions it can be shared.»[15]

[Rz 19] A longstanding Fair Information Practice Principle is data minimization. This notion of data minimization is extended to all aspects involving personal data. Applying safeguards to this data is also an essential component to privacy regulatory frameworks around the world. Appropriate safeguards are necessary to ensure that personal and confidential information in its custody and control are not compromised through inappropriate access, and that regulations within the jurisdictions in which they operate are not breached.

[Rz 20] Needless to say, organizations that can accomplish the transition to ABAC will develop a competitive advantage that will allow them to significantly reduce the risk profile of their access control regimen. By implementing a policy driven approach, they will be able to adhere to the tightening regulatory constraints on businesses, not to mention complying with data protection laws and they will simplify their software development and maintenance workloads. It is also fair to caution that venturing into the area of ABAC is not an easy migration for organizations with legacy access control mechanisms. It pre-supposes a mature identity and access management environment, in addition to a willingness to make core changes to an organization's application portfolio. That being said, if undertaken, the rewards of ABAC are significant — not only will the successful deployment of ABAC provide unprecedented control over access to protected resources, it will facilitate the establishment of business-focused, centralized policy management and a common, consistent access control model for applications across the enterprise. Such an approach not only addresses the complex challenges that Big Data applications pose to legal requirements involving personally identifying information by strengthening access to data through fine-grained access controls/permissions but also enhances legal compliance by taking a «default» approach to such assurances by directly embedding the access policy requirements into the system.

4.1 The use of roles in establishing entitlements

[Rz 21] **Role-based access control** (RBAC) has served us well and remains a positive and efficient way in which to automate the granting of permissions to users, based on their roles within an organization. When a person joins an organization, their access to systems and protected resources in the organization should be commensurate with their role(s). Furthermore, when one or more

[15] Nissenbaum, Helen. 2011. «A Contextual Approach To Privacy Online». Daedalus 140 (4): 32—48. doi:10.1162/daed_a_00113. See also Hart, Michael / Johnson, Rob / Stent, Amanda. 2013. «More Content — Less Control: Access Control In The Web 2.0». *Academic Paper, Stonybrook University (Sunysb)*. http://spin2013.cs.sunysb.edu//papers/cbac-w2sp07.pdf.

of their roles change, their access to systems should change accordingly. Most access control systems in use today are role-based. A typical access control solution is based on Active Directory groups; if a user's ID is in the group with access permissions to a protected resource, they will be granted access, otherwise they will be denied. Indeed role-based access control has served us well and remains

[Rz 22] Your description lacks precision. It seems that what you mean is that roles are used either to manually allocate access rights or that there is a hierarchical access right concept . an efficient way in which to automate the granting of permissions to users, based on their roles within an organization.

[Rz 23] When a person joins an organization, their access to systems and protected resources in the organization should be commensurate with their role(s). Furthermore, when one or more of their roles change, their access to systems should change accordingly. For instance, if a person is the Finance Manager they should have access to the financial management system at an appropriate level. If they then act as the Chief Finance Officer for two weeks while their manager is on vacation, they should be given access to the board of governors documentation for that period of time. If the manager is then transferred from Finance to HR, they should lose access to the Finance System and be given access to the HR system. This should happen automatically with human intervention limited to managing exceptions.

[Rz 24] RBAC is therefore an efficient way to manage staff access to protected resources. It becomes difficult, however, if there are too many roles. Over time, access control lists become bloated and becomes a source of inherent risks to the organization. Often, it is necessary for the organization to conduct a thorough assessment of roles assignment, which typically leads to a consolidation and reduction of roles to that which are essential.

4.1.1 RBAC Example

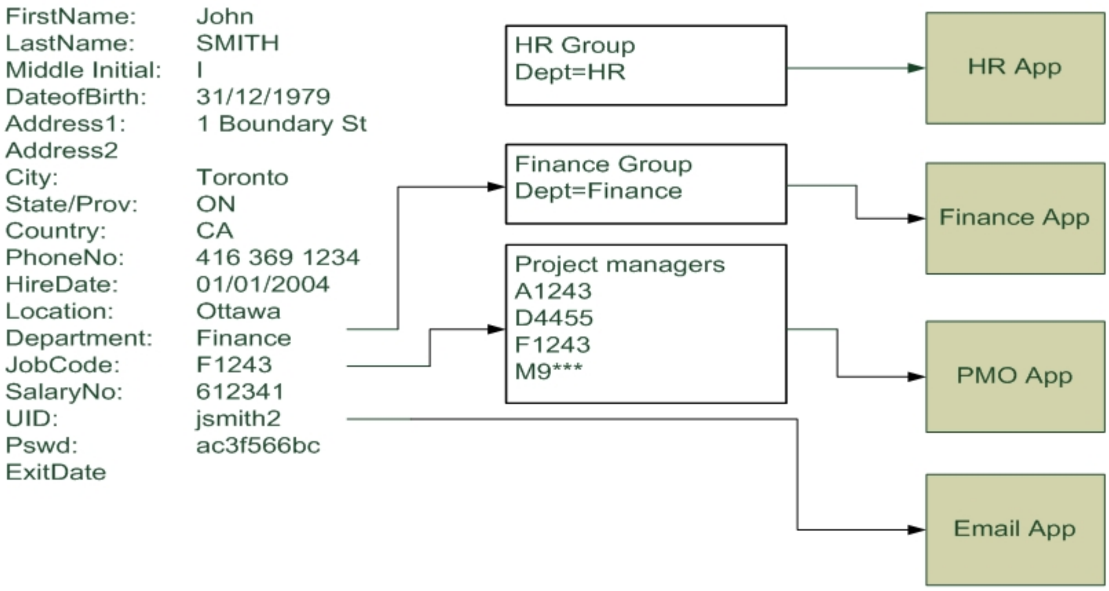
[Rz 25] There are many variants but all RBAC environments use a person's role information to grant access to applications. In

[Rz 26] Figure 1 below, John Smith is a Toronto-based staff member working in Ottawa. He gets access to the Finance application because he is in the Finance Group.

[Rz 27] This group can be populated automatically from the Department attribute in the organization's directory. The application needs to be configured to look at this attribute when a user logs in or the Finance application's access control list, or AD group needs to be populated from this attribute.

[Rz 28] The project manager's group is configured for JobCode F1243 so John also gets access to the Project Management Office application. The application's access control must look at the PMO group when authenticating users, and the appropriate job codes must be maintained in this group to ensure appropriate personnel get access to the application. All staff get provisioned into the email program but John does not get access to the HR application.

Figure 1 RBAC Example



4.2 The Attribute Approach

[Rz 29] In contrast to RBAC, attribute-based access control takes a very different

[Rz 30] Roles are attributes, too. What you say is rather that on the one hand ABAC is more general then RBAC as it includes all types of attributes, not just role attribute, and that there is a formalization in behind plus a mechanism to validate attributes.

[Rz 31] You should explain that more clearly. In addition, you should discuss the validation issues for attributes. approach to authorization. RBAC systems typically rely on group memberships (often referred to as «roles») as opposed to ABAC systems whereby access to protected resources is based on a user having specific attributes of which one's role may only be one such attribute (e.g. name, d.o.b., hire date, address, phone number, job title). In contrast to RBAC, attribute-based access control takes a very different approach to authorization. Access to protected resources is based on a user having specific attributes (e.g. name, dob., hire date, address, phone number, job title). This allows a much more fine-grained access control approach combining not only user attributes but other data, such as location (IP address or GPS) and time-of-day, to the access control decision. It allows unprecedented control of access to restricted resources based on fine-grained attributes evaluated at run-time. Rather than just using the role of a user to decide whether or not to grant them access to a system or protected resource, ABAC can combine multiple attributes to make a context-aware decision regarding individual requests for access.[16] Such systems do presume a mature identity and access management systems that collects and validates the desired

[16] See also discussion on «Adding Attributes to Role-Based Access Control» Kuhn, D. Richard / Coyne, Edward J. / Weil, Timothy R. 2010. «Adding Attributes To Role-Based Access Control». *Computer* 43 (6): 79—81. doi:10.1109/mc.2010.155. Merging the best features of RBAC and attribute-based systems can provide effective access control for distributed and rapidly changing applications.

attributes. Leveraging this rich store of data provides the organization a competitive advantage over their competitors lacking such a repository.

[Rz 32] The system rules based on these attributes can also incorporate contextual privacy requirements that organizations are required to comply with under data protection laws such as:

- Restrictions on outsourcing personal data (e.g. several jurisdictions expressly prohibit personal data disclosure to a jurisdiction without similar privacy legislation).
- Data minimization where only personal data that is required for the provision of the requested product or service should be collected by an organization. It is not permissible to collect data that «might be» useful at some point in the future.
- Purpose specification that only allows an organization to use personal data for the express purpose for which it was collected, and where any secondary use requires the consent of the individual.[17]

[Rz 33] While there are both open standard and proprietary approaches to support an attribute based access control system, the authors provide an industry open standard example known as the Extensible Access Control Markup Language (XACML)[18].

4.2.1 XACML - Data Loss Prevention

[Rz 34] Enterprises have legal, regulatory, and business reasons to protect their information, as exemplified by privacy, contracts, financial regulations, and export regulations. Organizations interpret those legal agreements, regulations, and business rules to form security and information protection policies, expressed in natural languages. Business policies and regulations are then instantiated as machine-enforceable access control policies. Most organizations employ a variety of security software tools to enforce access control policies and monitor compliance. In many cases, each tool must be configured independently of the others, leading to duplicative efforts and increased risk of inconsistent implementations.

[Rz 35] XACML-conformant access control systems provide scalable and consistent access control policy management, enforcement, and compliance for web services, web applications, and data objects in a variety of repositories. The XACML policy format and reference architecture can be extended to promote policy consistency and efficient administration in Data Loss Prevention (DLP) by adding in Policy Enforcement points into various DLP tools. DLP tools monitor «data-in-use» at endpoints (e.g., desktops, laptops, and mobile devices), «data-in-motion» on networks, and «data-at-rest» in storage systems. DLP tools enforce access control policies at these locations to prevent unauthorized access to and unintended disclosure of sensitive data. If DLP systems were to be standardized on the XACML policy format, enterprise policy authorities could use the same language to define access control policies for endpoints, networks, servers, applications, web services, and file repositories. The cost savings and improvements to security posture could be substantial. XACML policy format is suitable for and should be used to create, enforce, and exchange policies between different DLP systems. Subject information, including a rich set of me-

[17] Article 29 Data Protection Working Party. Statement of the WP29 on the impact of the development of Big Data on the protection of individuals with regard to the processing of their personal data in the EU. Adopted 16 September 2014. See also WP 29Opinion 03/2013 on Purpose Limitation.

[18] OASIS. eXtensible Access Control Markup Language (XACML) Version 3.0. January 2013. http://bit.ly/19x483Q.

tadata about subjects, will be expressed as subject attributes. Data objects and network resources will be expressed as resource attributes. Requests made by subjects and traffic operations will be expressed as action attributes.

[Rz 36] The use cases include:

Preventing sensitive data from being read/modified by unauthorized users
Preventing sensitive data from being emailed to unauthorized users
Preventing sensitive data from being transferred via web-mail
Preventing sensitive data from being copied from one computer to another
Preventing sensitive data from being transferred to removable media
Preventing sensitive data from being transferred to disallowed URLs

[Rz 37] Attribute-based access control holds significant promise for organizations with a good identity management environment, allowing access to protected resources to be based on more fine-grained attributes, based on context. That means that better overall security and in turn, data protection, is achieved and data governance is improved.[19]

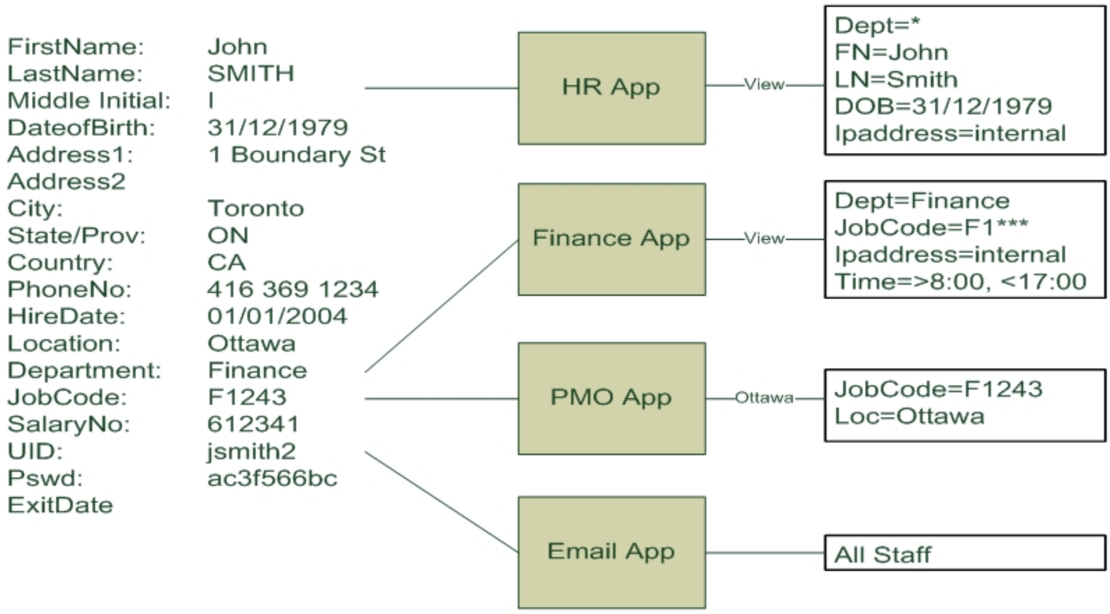
4.2.2 ABAC Example

[Rz 38] There are various approaches to ABAC[20] but they all rely on a rich data store of attributes generally aggregated from multiple authoritative sources. The core principle is to adhere to a published policy that determines the attributes required to gain access to protected resources.

[19] There are a number of use cases for ABAC such as research undertaken by staff at MITRE Corporation. 2015. Scalable Access Controls For Lineage. McLean Virginia. Accessed February 17. https://www.usenix.org/legacy/event/tapp09/tech/full_papers/rosenthal/rosenthal.pdf. Rosenthal, Arnon / Seligman, Len / Chapman, Adriane / Blaustein, Barbara.

[20] There is a an approach taken by e-Government Switzerland (e-CH 0107:IAM design principles) that defines the requirements, principles and rules that must be considered when deploying federated IAM solutions in eGovernment. The authors have been advised that this government standard describes an access mechanism that consists of a two-level granting of data access: the first level where widely accepted attributes are used; the second level where local attributes are used for fine-grained control. http://www.ech.ch/vechweb/page?p=dossier&documentNumber=eCH-0107&documentVersion=1.00.

Figure 2 ABAC Example



[Rz 39] In the above example, Figure 2, John will again gain access to the finance application because he works in the finance department, but only on a «view» basis because that's all that a «F1**» job code allows. He is also restricted to access during business hours. The access control system will allow him to access the PMO application, but only for the Ottawa projects.

[Rz 40] While John can't modify an HR system record, the use of ABAC allows for more fine-grained access to the application, allowing him to view his record. For HR staff, access can be segmented to roles: If an HR person has a recruiting job function, they should be permitted access to the recruiting sub-system; HR managers would get access as an administrator, and so on.

[Rz 41] The benefits of an ABAC approach are significant:

- access decisions are centrally managed via policies rather than by individual application managers, which means that they are applied consistently across the organization, and administered by business managers rather than controlled by IT personnel;
- software development is simplified by the incorporation of «policy enforcement point» code which externalizes the access decision to a «policy decision point;»
- decisions are made at runtime based on attributes that may be combined to form fine-grained decisions; changes in access status are immediately recognized when attributes are updated.

[Rz 42] From a business viewpoint, this translates to a vastly reduced risk profile; access rights will be modified as soon as the source system reflects the change. No longer will «old permissions» stay in the system when an attribute changes or a staff member leaves the organization's employ. Access rights are no longer based on access control lists that require manual intervention in order to be updated.

[Rz 43] It also means lower costs. Without manual intervention, there are no recurring costs for updating system access rights. Another benefit is that software development is less expensive, removing the cost of developing sophisticated access control logic for applications.

[Rz 44] There are several challenges, however, some of which are significant, when moving to an attribute-based access control environment. Firstly, a mature identity and access management environment is a prerequisite because a data model that has defined the «authoritative source» for attributes, and has efficient «source of truth» repositories, is an ABAC requirement. Authentication is a mission-critical service; it must be highly available and it cannot tolerate excessive network latency. This means that in some cases, a synchronization business model might be required whereby remote data stores are synchronized to a local directory or database.

[Rz 45] Policy Enforcement Points can also be a challenge, particularly if there are many legacy applications to integrate. Old systems with built-in access control logic will need to be modified to externalize the authorization function and take advantage of the more fine-grained control that ABAC can provide. Vendors can assist this development with code and APIs that facilitate entitlement management.

[Rz 46] Policy Decision Point management can be complicated for distributed organizations or hybrid situations in which some applications are in the Cloud. While policy management must remain centralized, the decision points will need to be distributed and a mechanism to keep them current will be required.

[Rz 47] Finally, Policy Administration must be addressed. With the movement of policy management from IT to business, there is a need for business units to understand their requirements and have a capability to encode their requirements in policies. In some organizations this may be hard to achieve and a centralized policy management facility, servicing multiple business units, will be required.

[Rz 48] It is also vital that these ABAC capabilities be developed and deployed using a common foundation of concepts and functional requirements to ensure the greatest level of interoperability possible.[21]

[Rz 49] Nonetheless, in the U.S., in 2009 the Federal Chief Information Officers council published the federal identity, credential and access management (FICAM) roadmap and implementation document which provides guidance to federal organizations to evolve their logical access control architectures to include the evaluation of attributes as a way to enable access within and between organizations across the federal enterprise. Ultimately, FICAM called out ABAC as a recommended access control model.[22]

4.3 Industry Examples and Use Cases

[Rz 50] Attribute based Access Control (Policy based Authorisation) is most applicable to those industries which are outlined below:

- Highly distributed with remote operating entities, perhaps in a subsidiary and parent environment, where these subsidiaries need to have on-line access to applications and services and where the principals are wanting to regulate subsidiary access and authorisation based

[21] Hu, Vincent C. / Ferraiolo, David / Kuhn, Rick / Schnitzer, Adam / Sandlin, Kenneth / Miller, Robert / Scarfone, Karen. Guide to Attribute Based Access Control (ABAC) Definition and Considerations. US Department of Commerce, NIST Special Publication 800-162 Guide to Attribute Based Access Control (ABAC): Definitions and Considerations. http://1.usa.gov/1aJynBD.

[22] Ibid.

upon a set of policies. Examples of these industries are:

  - Financial Services sector
  - Insurance sector
  - Travel Agents
  - Airlines
  - Telecommunications companies in delivery of segmented services to clients

- Organisations that are mandated by government or regulatory policy to control what information is provided or what goods may be exported. Examples of these organisations are:
  - Aerospace Manufacturers
  - Defence Manufacturers
  - Nuclear Energy Organisations
  - Chemical Manufacturers
  - Oil & Gas companies
  - Exporters
  - Customs Authorities
  - Barrier Control organisations such as Meat Export, or Fruit Export organisations (Australian Quarantine & Inspection Service)
- Organisations that have a need for strong protection of their Intellectual Property and wish to protect that IP by the implementation of Policy Based Data Loss Prevention services. Examples would be:
  - Pharmaceuticals companies
  - Research & Development Organisations
- Organisations that have a requirement involving sensitive health data with associated consent management requirements. Examples in these cases would be:
  - Hospitals
  - Primary care (e.g. physicians)
  - Secondary Healthcare Organisations
  - Tertiary Healthcare Organisations
  - Pharmacies
- Organisations that have a requirement to control access to Geospatial data and imagery and to share it using ABAC. This Imagery might be either highly classified or of more general nature for instance in disaster management; Examples of users would be:
  - Emergency Management & First Responders
  - Intelligence and Defence organisations such as the US defence Information Systems Agency, the German Federal Department of Defence and Australian Department of Defence, the US Geospatial Intelligence Organisation, the US National Security Agency (on Big Data projects such as the Ozone Widget Program) and agencies such as the Canadian, United Kingdom, Australian and NATO Defence Organisations.
  - Statre & Federal Police Forces
- Government organisations that have a requirement to protect the data in their custody and control, which they share or distribute through the use of security classifications systems, where ABAC can be applied based upon the combination of document or Big Data security, the person seeking to have access to that data, any obligations required such as redacting portions of the document and other resource attributes and factors such as location, time of

day, environment that need to be considered by the ABAC Policy servers when making the decision to allow or deny access.

[Rz 51] Attribute based Access control using either XACML or its Geospatial extension Geo-XACML has applicability to all of the above organisations. For many of them the OASIS XACML Committees has developed specific implementation profiles for instance for Healthcare, IP Protection, Government mandated controls on Exports etc. Other industry sectors are also incorporating ABAC Policy based Authorisations capabilities into their standards, an example being the Secure Messaging community and Network Management community. Big Data or access to it and control of it is another significant community where Policy based Authorisation services lend themselves to using ABAC rather than RBAC.

[Rz 52] The following use cases help to better understand the capabilities of an ABAC approach and its applicability to specific organizational needs.

4.3.1 Government, Security and Classified Intelligence Documents

[Rz 53] A major use of ABAC in Government is for security labelling, classifying both documents and emails to effectively control access to a particular document, but to also block evidence that such a document exists. For instance, a particular document might have a Meta Tag on it restricting access to USEO (United States Eyes Only) where only US Citizens are permitted to access the document and would even know that the document even exists on a particular file store. This use case combines attributes about the person requesting access and the resource they are attempting to access.

[Rz 54] By extension, if the Meta Tag classification on a document is AUSCANUKNZUS Eyes Only and there is a segment in the document which has a Top Secret addenda attached, then anyone with a Secret clearance may only access the Secret Portion of the document and only if they are nationals of Australia, NZ, Canada, UK or US. So for instance other NATO citizens with a Nationality of say Germany (GE) would not be allowed access to it even though they have Top Secret clearance as they would fail on the Nationality. ABAC then has further granular capability of restricting access to this Top Secret addenda for nationals of Australia, NZ, Canada, UK or US based upon Caveats and Code-words which would be interpreted as part of the Policy evaluation. It is through an obligation that this specific restriction can be applied. In other words, if access to the complete document is requested, the Policy server can instruct the Document Management system to only release the document once the Top Secret portion is redacted or deny access to it entirely. An Action may be a follow up obligation such as notifying a security officer that access has been requested.

4.3.2 Health Care

[Rz 55] The health care sector presents a multitude of use cases for which ABAC applies. In the U.S., the Affordable Care Act and Accountable Care Organizations stretch across the continuum of health care and are taxing the capability of role-based access controls. For example, cardiologists who practice in the General Hospital Electrophysiology clinic may be allowed to view a person's preliminary and final ECG while they are on the hospital network. Other physicians are not allowed to see the preliminary reports - only the final interpretation. Researchers, however, are only allowed to see a de-identified or redacted version of the patient's interpreted ECG. And no-one is permitted to see the ECG outside of the General Hospital network. This complexity can only be managed through attribute-based access controls. Standards such as XACML provide a flexible and dynamic access control policy framework.'[23]

[Rz 56] ABAC is also important from a patient perspective. In order to effectively provide consent management to their electronic health records, an attribute-based solution is ideal to provide the level of granularity required to allow patients to manage their own sensitive information.

4.3.3 Telecommunications Carriers

[Rz 57] An interesting development in the use of ABAC has been to enable subscriber-managed access controls for broadband and mobile services. Usually, telecommunication providers create a customer account for each type of service offered (e.g. mobile, broadband, pay TV, games) alongside its own customer engagement model and its own subscriber registration process. They have attempted to deploy a simple roles-based or group-based access control (RBAC) mechanism to manage a customer's ability to turn on applications or access content resources by the central administration team. Whenever a new service offering is introduced, a new role and group must be created and the specific customer/subscriber provisioned. Over time, the creation and deprecation of applications and services results in the inevitable propagation of roles and groups, making administration of the customer provisioning environment overwhelming.

[Rz 58] With ABAC, an access control account may be created for a particular customer/subscriber, empowering the customer/subscriber to provision one's own users e.g. children having access to children's games but not Internet or only for two hours a day, the family only having access to adult content on Pay TV after 11pm, perhaps financial limits on mobile calls. The value is that this activity will be driven by policy and authorisation services externalised from the application and not embedded in the application itself. The policy engine combining these rights and attributes uses the XACML Combining algorithm. Moreover, rather than provisioning the device or subscriber at the time of contract, the telecommunications provider has the ability to undertake policy based provisioning, thus dynamically provisioning a subscriber's access to a given service or content at the time of the request. This capability enables a much more dynamic service model.

4.3.4 Insurance

[Rz 59] The Insurance industry also has significant benefit in adopting an ABAC approach. There are multiple players in the insurance supply chain from agents to underwriters to claims adjusters to clients. Each participant has a different information requirement and confidentiality level which is difficult to manage via an RBAC environment. Additionally, the centralized access policy management functions of an ABAC environment make the administration of the access control task significantly more manageable for the insurance industry participants.

[23] PHEMI. 2014. Balancing Access To Information While Preserving Privacy, Security And Governance In An Era Of Big Data. White Paper. Vancouver, B.C.: PHEMI. http://www.phemi.com/wp-content/uploads/2014/06/White-Paper-Privacy-Security-Governance-in-the-Era-of-Big-Data.pdf.

4.3.5 Airlines

[Rz 60] Airline operations have a strict security regimen that must be observed. ABAC has clear advantage here since it can provide fine-grained access control. For instance, airline personnel typically need to access secure areas in an airport and are typically regulated by a government body that must ensure appropriate training and oversight are applied to staff and contractors. As such, it is necessary to monitor personnel with air-side access at airports and monitor the number of access instances in a specific period. Access in some cases is geographic or may be time-of-day dependent. ABAC has the capability to manage such access and also to facilitate management of the policies that determine the access permissions to be granted. There is also a public-facing requirement in managing a client's access based on loyalty programs, travel agent status or airline partner relationship. ABAC may be employed to quickly permit or deny a person's access rights via a central policy management tool.

5 Conclusion

[Rz 61] The emergence of Big Data and the burgeoning business associated with the Internet of Things, introduce not only unprecedented opportunities, but at the same time, considerable challenges to security and privacy. An organization's legal requirements to safeguard personal information from unauthorized access and use rely on a number of automated systems and controls. A strong approach to reducing the risks to security and privacy and enhanced legal compliance through advanced attribute-based access controls, ABAC has significant potential and should be considered by organizations with an identity and access management environment. It allows an unprecedented amount of flexibility and security while promoting information sharing between diverse and disparate organizations — consistent with a Privacy by Design framework. In the context of Big Data and cloud computing, granular access control gives data managers a scalpel instead of a sword[24] to share data as much as possible, without compromising privacy or security. Even if an organization is not considering adoption of ABAC in the immediate future, it may be positioned in the organization's IT roadmap to ensure that any IT infrastructure extensions are aligned with an ABAC approach. Any organization developing or acquiring new applications should consider ABAC to ensure that the new deployments are «future proofed» by externalizing authorization management.

6 References

- European Commission. 2014. *Progress On EU Data Protection Reform Now Irreversible Following European Parliament Vote.* http://europa.eu/rapid/press-release_MEMO-14-186_en.htm.
- Federal Trade Commission (FTC). 2012. *Protecting Consumer Privacy In An Era Of Rapid Change: Recommendations For Businesses And Policymakers.* http://www.ftc.gov/news-events/press-releases/2012/03/ftc-issues-final-commission-report-protecting-consumer-privacy.

[24] Cloud Security Alliance. Top Ten Big Data Security and Privacy Challenges. November 2012.

- «Privacy By Design Resolution». 2010. In *32Nd International Conference Of Data Protection And Privacy Commissioners*. Jerusalem.
- *Ten Tips For Reducing The Likelihood Of A Privacy Breach (Fact Sheet)*. 2014. Ottawa, Ontario: Office of the Privacy Commissioner, Canada. https://www.priv.gc.ca/resource/fs-fi/02_05_d_60_tips_e.asp.

7 Privacy and Big Data Institute, Ryerson University

[Rz 62] In the knowledge economy, the world is increasingly being mapped in vast sets of data points. With Big Data, information may be viewed as currency. Big Data is opening opportunities in business, medicine, education, city planning and development — virtually every sector of the economy.

[Rz 63] The Privacy and Big Data Institute (PBD Institute) at Ryerson University, led by Dr. Ann Cavoukian, is an academic-industry partnership that will be developing the methodologies relating to the expanding world of Big Data, in a manner that ensures the presence of privacy-preserving measures embedded into the data analytics. We will be leading the research and development of Big Data applications by ensuring they are proactively embedded in a Privacy by Design framework. Ours is the only Big Data Institute in the world that has branded itself with Privacy.

[Rz 64] Housed in Ryerson's Faculty of Science, the Privacy and Big Data Institute brings the cross-discipline synergy of the university setting to data-driven research and innovation, and becomes the immediate training ground for next-generation talent.

[Rz 65] Ryerson researchers in mathematics, science, engineering, commerce and more are studying Big Data in ways that affect business healthcare, public policy and everyday life. The institute will support this research, foster the growth of privacy-enhanced Big Data expertise and provide an educational platform through publications, panels, seminars, undergraduate and graduate course, continuing education programs and more.

[Rz 66] The Privacy and Big Data Institute dovetails with the existing innovation ecosystem at Ryerson, encouraging the use and development of Privacy by Design applications and technologies.

[Rz 67] The result will be a dynamic, multi-disciplinary, integrated experience at Ryerson powered by privacy-enhanced Big Data analytics and applications: tools, training, commercialization and adoption. The Privacy and Big Data Institute will direct those energies to address real-world issues.

http://www.ryerson.ca/pbdi/

Ann Cavoukian, Ph.D., Executive Director, Privacy and Big Data Institute, Ryerson University, Toronto, Canada. Ann.Cavoukian@ryerson.ca.

Michelle Chibba, M.A., Independent Privacy/Strategic Policy Expert, Toronto, Canada.

Graham Williams, BSC, MBA, Global Analyst, KuppingerCole and Director, KuppingerCole (Asia Pacific) Pte Ltd. gw@kuppingercole.com.

Andrew Ferguson, Industry Specialist in Identity and Access Management and Director, KuppingerCole (Asia Pacific) Pte Ltd. af@kuppingercole.com.