White paper

Big Data Analytics and privacy & data protection

Your business technologists. **Powering progress**

Atos Consulting

Abstract

The rise of Big Data and Data Analytics provides great opportunities for organizations to realize new ways of doing business and will give a significant increase in added value. Organizations that do not 'catch the Big Data Analytics train' run the risk of falling behind. According to the European Commission[1] the value of European citizens' personal data has the potential to grow to nearly € 1 trillion a year by 2020.

Lawmakers around the globe, responding to many well-publicized situations and concerns expressed by citizens, have become more and more focused and concerned with privacy matters. An important development in this respect is the current reform of the European Union's (EU) data protection legislation framework which revolves around the adoption of a new 'General Data Protection Regulation[2]' (hereinafter referred to as 'the Regulation'). The Regulation will replace the current European Data Protection Directive[3] (EU Directive 95/46/EC) and is expected to be adopted by the end of 2015. It will bring major changes to data protection legislation in Europe (e.g. a significant increase of the potential sanctions incurred for non-compliance with the legislation: up to € 100 million per event, or 2 to 5% of the annual turnover).

Uncertainty about the data protection rules are causing organizations to miss out on business opportunities and innovations with Big Data and Data Analytics. How can organizations significantly improve their efficiency and offerings with Big Data Analytics while implementing the relevant privacy & data protection principles and rules? This paper provides the answer to this question.

Privacy & data protection

White Paper on how organizations can significantly improve their efficiency and offerings with Big Data Analytics while implementing the relevant privacy & data protection principles and rules.

Contents

Introduction 3

Terminologies 4
- What is Big Data?
- What is Big Data Analytics?
  - Becoming an Analytics Driven Organization
- What is privacy and personal data?
  - Definition of privacy
  - European Awareness
  - Definition of personal data
  - Special categories
  - Legal aspects and privacy

Privacy & Personal Data challenges 8

General Challenges
- Specific Challenges
  - OECD privacy principle 1 Collection Limitation Principle 8
  - OECD privacy principle 2 Data Quality 9
  - OECD privacy principle 3 Purpose Specification 10
  - OECD privacy principle 4 Use Limitatiom 10
  - OECD privacy principle 5 Security Safeguards 11
  - OECD privacy principle 6 Openness 12
  - OECD privacy principle 7 Individual Participation 13
  - OECD privacy principle 8 Accountability 13

NIST and Big Data 15

Big Data Interoperability Framework
- Security & Privacy Subgroup
- Mapping challenges to NBDRA

Conclusion 20

Colophon 21

[1] European Commission (April 2015). The EU Data Protection Reform and Big Data Factsheet. Brussels, Belgium, European Commission Directorate-General for Justice and Consumers.
[2] European Commission (25 January 2012 - COM(2012) 11 final). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Brussels, Belgium, European Commission Directorate-General for Justice and Consumers.
[3] Official Journal of the European Union (23 November 1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Luxembourg, Luxembourg, EUR-Lex.

Introduction

The rise of Big Data and Data Analytics provides great opportunities for organizations to realize new ways of doing business and will give a significant increase in added value. Organizations that do not 'catch the Big Data Analytics train' run the risk of falling behind. According to the European Commission[4] the value of European citizens' personal data has the potential to grow to nearly €1 trillion a year by 2020.

"The value of European citizens' personal has the potential to grow to nearly €1 trillion annually by 2020"

Lawmakers around the globe, responding to many well-publicized situations and concerns expressed by citizens, have become more and more focused and concerned with privacy matters. An important development in this respect is the current reform of the European Union's (EU) data protection legislation framework which revolves around the adoption of a new 'General Data Protection Regulation[5]' (hereinafter referred to as 'the Regulation'). The Regulation will replace the current European Data Protection Directive (EU Directive 95/46/EC[6]) and is expected to be adopted by the end of 2015. It will bring major changes to data protection legislation in Europe. Amongst the major evolutions expected from the new legislation and relevant to this paper will be:

- The scope of the legislation which will now make European principles unquestionably applicable to data controllers established outside the EU but targeting residents in the EU;
- The principle of accountability for data controllers, i.e. the obligation to be able to demonstrate, at all times, compliance with applicable data protection legislation (through documentation, processes, etc.);
- The express recognition of the 'right to be forgotten';
- The requirement of 'privacy by design', i.e. taking into consideration data protection considerations when developing a new product or service;
- The generalization of Privacy Impact Assessments (PIA) for potentially privacy & data protection risky processes;
- The significant increase of the potential sanctions incurred for non-compliance with the legislation: up to € 100 million per event, or 2 to 5% of the annual turnover.

Uncertainty about the data protection rules are causing organizations to miss out on business opportunities and innovations with Big Data and Data Analytics. How can organizations significantly improve their efficiency and offerings with Big Data Analytics while implementing the relevant privacy & data protection principles and rules?

This White Paper addresses this question by first presenting the notions of Big Data, Big Data Analytics and privacy & personal data terminologies (Chapter 4). We then go on to describe the key privacy & data protection principles and their implications (Chapter 5). The challenges created by privacy & data protection to the implementation of Big Data Analytics and our recommendations to design 'privacy-friendly' Big Data Analytics solutions are addressed also in chapter 5. Furthermore, we introduce the recently published Institute of Standards and Technology (NIST) Big Data Reference Architecture (NBDRA) standard and present our privacy & data protection challenges in the NBDRA (chapter 6). Finally, the answer to the question, 'how can organizations significantly improve their efficiency and offerings with Big Data Analytics while implementing the relevant privacy & data protection principles and rules' is given in chapter 7.

[4] European Commission (April 2015). The EU Data Protection Reform and Big Data Factsheet. Brussels, Belgium, European Commission Directorate-General for Justice and Consumers.
[5] European Commission (25 January 2012 - COM(2012) 11 final). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Brussels, Belgium, European Commission Directorate-General for Justice and Consumers.
[6] Official Journal of the European Union (23 November 1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Luxembourg, Luxembourg, EUR-Lex.

Terminologies

What is Big Data?

It has become unnecessary to argue the fact that the amount of personal data has exploded over the past few years and continues to grow exponentially. The success of companies like Google, Facebook, Youtube, Twitter and LinkedIn, the development of the Internet of Things (IoT) and the success of ever more powerful smartphones have created the need to collect, store, transfer, process and analyze extremely large volumes of (personal) data.

Prior to the search engine and social media revolutions, organizations also processed large volumes of data[7], but unlike traditional data(base) management systems, Big Data(base) management systems are now able to store, process and analyze 'new' and different types of data from a variety of sources in a way that was not possible before. The emergence of new technologies has also reduced costs for organizations to conduct these activities.

On the basis of these elements, Big Data can be described by the following data characteristics[8]: extremely large volumes, great variety of data and enormously high velocity of generation. Big Data systems must therefore be capable of handling these three 'V's.

Three 'V's' of Big Data:
1. Extremely large volume of data,
2. Enormously high velocity of data,
3. Great variety of data

So the purpose of Big Data systems is to be able to 'dig out' value in data that was initially not exploited because of the lack of fully handling the three 'V's' characteristics. The concept behind Big Data is thus to identify trends, characteristics or other elements from information from different sources which could not initially be efficiently analyzed together.

What is Big Data Analytics?

Data Analytics is more than just analyzing a large amount of data. It's about creating unique and actionable insights by combining the right data types and sources for the right purpose. Therefore organizations should have clear strategic business goals in mind before putting Data Analytics to work. The questions organizations should ask themselves are e.g.:

- Clear targets: "What is the desired outcome with Data Analytics to improve the customer experience?"
- Data sources and technology: "What kind of data and technology is already in-house? And "how do we get access to and use new data sources and technology?"

Having clear targets and the needed data and technology will not automatically create value out of Data Analytics. For this to happen, organizations should create an analytic environment which drives the analytical spirits of employees. The organization should further be aware of the needed analytical resources and envision how to develop them. Finally, the organization should be aware of risks, compliancy- and security policies that go along with using customer data.

Becoming an Analytics Driven Organization

Recently, we released our white Paper on 'the Analytics Driven Organization[9]'. In this paper we defined our vision on what it takes to be a truly leading Analytics Driven Organization, via a model of nine different dimensions. The nine dimensions are[10]:

1. Analytics leadership;
2. Analytics strategy;
3. Analytical culture;
4. Analytics capability and governance;
5. Analytics skills and competences;
6. Data strategy;
7. Analytics technology strategy;
8. Analytics performance and risk management;
9. Analytics security and compliance.

In brief, the Analytics Driven Organization paper advocates that organizations which use tools such as Hadoop aren't necessary leading in 'Big Data Analytics'. To be or become a leading Big Data Analytics player, an organization should take into account developing all nine dimensions.

[7] Irrespective of (public) sources.
[8] The term 'Volume, Velocity and Variety' was first introduced by Laney Doug (6 February 2001). APPLICATION DELIVERY STRATEGIES - 3D Data Management, Controlling Data Volume, Velocity and Variety. Stamford, USA, META Group Inc.
[9] Fichtinger, R. and Mallison, N. (July 2015). Whitepaper – The Analytics Driven Organization. Utrecht, The Netherlands, Atos Consulting.
[10] The nine dimensions are explained in the whitepaper on 'The Analytics Driven Organization' (see reference 9).

What is privacy and personal data?

Definition of privacy

A sound and universal definition for privacy is difficult to find. The earliest definition of privacy was given by Warren & Brandeis[11]. They defined privacy as 'the right to be left alone'. In this paper we are using the Warren & Brandeis definition in relation to personal data protection (see also next paragraph in this chapter).

"Privacy is 'the right to be left alone'"

European Awareness

The European awareness and attention to privacy and data protection first materialized in the 'European Convention for the Protection of Human Rights and Fundamental Freedoms"[12] in 1950. This convention also established the European Court of Human Rights (ECtHR) in Strasbourg. The ECtHR hears any citizen who feels that his or her rights under the Convention have been violated by a Member State. The first international and legally binding instrument specific to personal data protection (and privacy) was the Council of Europe's Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, adopted in 1981[13] (better known as "Convention 108"). That same year the Organization for Economic Co-operation and Development (OECD) published its 'Guidelines Governing the Protection of Privacy and Trans-Border Flows of Personal Data14', which set out seven (later eight) principles for protection of personal data (see textbox OECD).

The OECD data protection principles

1. Collection Limitation (limits on collecting of personal data)
2. Data Quality (Personal data must be relevant to the purposes, must be accurate, complete and kept up-to-date)
3. Purpose Specification (collection purposes must be specified and limited to the fulfillment of the purpose)
4. Use Limitation (Personal data must not be disclosed, made available or otherwise used than other than specified and with the consent of the individual (data subject) or by the authority of law)
5. Security Safeguards (Personal data must be protected)
6. Openness (Organization must have general policy of openness about developments, practices and policies with respect to personal data)
7. Individual Participation (individual must have the right to control the personal data relating to him)
8. Accountability (the data controller (a person or body that determines the purpose, condition and means of processing personal data) is accountable to comply with the seven principles above).

[11] Warren, S.D. and Brandeis, L. D. (15 December 1890 - pp. 193-220). The Right to Privacy. Boston, USA, The Harvard Law Review Association. .
[12] Council of Europe (4 November 1950). Convention for the Protection of Human Rights and Fundamental Freedoms as amended by Protocols No. 11 and No. 14. Rome, Italy, Steering Committee for Human Rights, European Court of Human Rights.
[13] Council of Europe (28 January 1981). Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108). Strasbourg, Germany.
[14] Organization for Economic Co-operation and Development (11 July 2013 - C(80)58/FINAL, as amended on 11 July 2013 by C(2013)79). Recommendation of the Council concerning Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data. Paris, France.

Definition of personal data

The notion of personal data (or personally identifiable information, depending on the country specific jurisdictions) has generally received a relatively similar definition around the globe. In this respect, it is safe to say that the definition adopted in the European legislation that it is 'information relating to an identified or identifiable natural person' (data subject), is generally accepted (even though some jurisdictions such as Switzerland go beyond that by including data relating to corporate bodies).

Data is considered personal when someone is able to connect information to a specific person, even when the person or entity that is holding the personal data cannot make the connection directly (e.g. name, address, e-mail), but has or may have access to information allowing such identification (e.g. through telephone numbers, credit card numbers, license plate numbers, etc.).

Special categories

Certain categories of data are by nature very intimate to a person or may be used in a way which may be detrimental to the person concerned in the absence of any legal justification (i.e. discrimination). Such data has always been regarded as requiring particular attention to avoid excessive and/or unfair exploitation. It is again relatively widely recognized that such special categories of personal data (or 'sensitive personal data') include, as per the EU legislation:

- personal data revealing racial or ethnic origin;
- personal data revealing political opinions, religious or other beliefs and;
- personal data concerning health or sexual life.

Due to this specific care, which is also subject to local history, culture and principles, it is generally prohibited to process such categories of data unless adequate measures are in place to guarantee adequate protection for this data and circumstances require it (e.g. required under a legal obligation and/or with the explicit consent of the data subject).

Legal aspects and privacy

The principles relating to the fashion in which data should be processed are set out not only in legislation (i.e. the European data protection legal framework) but also in what can be regarded as 'soft law' (e.g. statements, principles, codes of conduct, codes of practice, ect.).
As previously mentioned, the OECD has adopted privacy principles but other relatively similar principles can be found in the International privacy framework standard 'ISO/IEC 29100[15]' and the 'Generally Accepted Privacy Principles[16]' (GAPP) which have been developed by the American Institute of Certified Public Accountants (AICPA) and Canadian Institute of Chartered Accountants (CICA). The GAPP provide best practice controls to protect personal data and to help organizations with defining and implementing their personal data protection programs.

Table 1[17] below establishes a form of correspondence table between the principles set out in these various sources, allowing us to witness the similarities between them regardless of the way in which they have been formulated.

[15] The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) (15 December 2011). ISO/IEC 29100:2011, Information technology – Security techniques – Privacy framework. Geneva, Switserland, ISO/IEC.
[16] American Institute of Certified Public Accountants and Canadian Institute of Chartered Accountants (August 2009). Generally Accepted Privacy Principles, business versions. USA.
[17] Based on the initial mapping done by Blanchard, C. (29 October 2013). Presentation - data privacy and protecting personal data. Session 142 Information Security and Risk Management conference. Milwaukee, USA, ISACA

Table 1: Relation between OECD privacy principles, EU Directive, ISO/IEC 29100 and GAPP

OECD privacy priciples	EU Directive personal data requirements	Generally Accepted Privacy Principles	ISO / IEC 29100 privacy principles
1. Collection Limitation	- Consent (art. 2) - Data Quality (art. 6) - Legitimacy of processing (art. 7) - Exemption and restriction (art. 13)	- Collection - Choice and consent	- Consent and choice - Collection limitation
2. Data Quality	- Data quality (art. 6)	- Data quality	- Accuracy and quality
3. Purpose Specification	- Information to be given to data subject (art. 10 and 11) - Data quality (art. 6) - Lawful processing of personal data (art. 5 and 7)	- Notice	- Purpose legitimacy and specification
4. Use Limitation	- Data quality (art. 6) - Data processing limitation (art. 7) - Data subject rights to object (art. 14) - Transfer of data to Third countries (art. 25 and 26)	- Use, retention and disposal - Disclosure to third parties	- Data minimization - Use, retention and disclosure limitation
5. Security Safeguards	- Security of processing (art. 16 and 17)	- Security of personal data	- Information security
6. Openness	- Information to be given to data subject (art. 10 and 11)	- Notice	- Openness, transparency and notice
7. Individual Participation	- 'Right of access' (art. 12) - Judicial remedies, liability and sanctions (art. 22, 23 and 24) - Supervisory Authority (art. 28)	- Monitoring and enforcement - Access	- Individual participation and access
8. Accountability	- Security of processing (art. 16 and 17) - Notification (art. 18, 19 and 20)	- Managament	- Accountability - Privacy compliance

Privacy & Personal Data Challenges

In the previous chapter we introduced the OECD privacy principles and mapped those principles to the EU Directive requirements, the ISO/IEC 29100 privacy principles and the GAPP principles. In this chapter we explain the OECD privacy principles in more detail to further clarify how they present challenges for Big Data Analytics. We also provide the applicable GAPP privacy objectives associated with the corresponding OECD data protection principles. Finally, we discuss one example control to design privacy & data protection friendly Big Data Analytic solutions.

General Challenges

By definition, Big Data Analytics are faced with important privacy & personal data issues. Indeed, the very purpose of Big Data Analytics is to exploit extremely diverse data. This without necessarily having defined the objective and purpose for which the data are collected, analyzed, stored and processed (to information). Big Data has fundamentally changed the paradigm of data processing, challenging data protection specialists in their approach to the way data protection principles should or could be applied to this new paradigm. The fact that Big Data calls for the processing of vast amounts of data ('volume') naturally raises issues concerning the quality of that data. The fact that Big Data relies on diverse sources ('variety'), questions the reliability of certain results which can lead to inappropriate profiling or discrimination as well as the level of confidentiality of data. The fact that systems can now process vast volumes of data extremely quickly ('velocity') questions the possibility to 'de-identify' or anonymize data efficiently when cross-references are so numerous. Combine these elements and further questions arise regarding the possibility of efficiently implementing Big Data Analytics. Also, for instance, the compliancy with the legitimacy of the processing and the purposes for which it has been implemented.

Specific Challenges

In this paragraph we provide practical examples of designing privacy & personal data compliant and user friendly Big Data Analytic solutions. This is done by addressing the applicable GAPP privacy objectives associated with the corresponding OECD data protection principles, followed by the selected GAPP control (measure).

OECD privacy principle 1: Collection Limitation Principle

There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.

Big Data Analytics privacy & data protection challenge
This principle possibly has the greatest impact on Big Data Analytics. The essence of Big Data Analytics is to collect, store, transfer, process and analyze large volumes of extremely varied (personal) data at high velocity which do not necessarily have any direct correlation at the time of processing,. Therefore, Big Data Analytics is focused on data maximization and seeking promising correlations when privacy & personal data protection principles call for data minimization.

The variety of sources of the data also raises questions on ensuring that all data included in a Big Data Analytics data set has been collected legitimately and on liability for ensuring such legitimacy. In addition, the issue of data subject knowledge or consent of the processing may also be problematic (as addressed in the "Openness" principle below): while they may have been informed of certain processes related to their data, how can it be ensured that adequate and complete information has been provided for all aspects of the processing? From a legal standpoint, the challenge posed by Big Data to this principle, is to ensure that the data processed has been or is obtained in a legitimate fashion, i.e. in compliance with applicable laws and without deceiving the data subjects. One of the main issues in order to achieve this objective relates to the origin of the data. Indeed, the entities implementing such practices will often combine databases containing data they have not obtained directly and therefore they have no direct control over how the data was initially collected.

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objectives are applicable for this challenge.

- [GAPP 3.0] The entity describes the choices available to the individual and obtains implicit or explicit consent with respect to the collection, use, and disclosure of personal information.
- [GAPP 4.0] The entity collects personal information only for the purposes identified in the notice.

Measure to address the privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 3.2.1] Implicit or explicit consent is obtained from the individual at or before the time personal information is collected or soon after. The individual's preferences expressed in his or her consent are confirmed and implemented.

Atos' Vision
In its approach to Big Data Analytics, Atos aims at ensuring that the data it receives has been collected adequately. From a legal standpoint, the solution which has been identified is twofold. Firstly, Atos requires that its agreements with data providers includes adequate provisions whereby the provider guarantees that it has collected the data adequately and has obtained the necessary consent from the data subjects to share the information with Atos for analytics purposes. Secondly, Atos reviews the terms of the information notice provided to the data subjects to ensure that it has effectively the right to process the data in question.

[18] See the AICPA website (http://www.aicpa.org) for the GAPP standard with the complete overview of the GAPP controls.

In practice for principle 1: Collection Limitation
The entity:

- Obtains and documents an individual's consent in a timely manner (that is, at or before the time personal information is collected or soon after).
- confirms an individual's preferences (in writing or electronically).
- documents and manages changes to an individual's preferences.
- ensures that an individual's preferences are implemented in a timely fashion.
- addresses conflicts in the records about an individual's preferences by providing a process for users to notify and challenge a vendor's interpretation of their contact preferences.
- ensures that the use of personal information, throughout the entity and by third parties, is in accordance with an individual's preferences.

OECD privacy principle 2: Data Quality

Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.

Big Data Analytics privacy & data protection challenge
Unlike traditional data(base) management systems, data modelling in Big Data Analytics systems is not always necessary. Some approaches simply collect and store the data without processing and/or without a specific predefined purpose. This also creates a difficulty with regards to the requirement that data only be retained for a period of time which is not excessive in regards to the purpose for which it is collected. In the absence of a defined purpose (as addressed below) the matter of data retention becomes even more important. Data which may be regarded as 'stale[19]' in many instances can be regarded in Big Data Analytics as providing valuable information.

Therefore, the integrity of the stored (personal) data is not validated and potential errors in the data set may remain. Once the data is (being) processed the errors may be detected. Some of the correlations and results of Big Data Analytics may therefore result in inaccurate results which can have negative or discriminatory effects on the data subjects concerned. This principle therefore necessitates that Big Data Analytics operators define clear rules regarding the data they wish or intend to process. Failure to clearly define the necessity or relevance of potentially 'stale' data will expose Big Data Analytics operators to non-compliance and corresponding sanctions.

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objective is applicable for this challenge.

- [GAPP 9.0] The entity shall maintain accurate, complete, and relevant personal information for the purposes identified in the notice.

Measure to address the privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 9.2.1] Personal information is accurate and complete for the purposes for which it is to be used.

Atos' Vision
In order to address this matter, Atos recommends putting in place a full data mapping, defining a clear view of what data is being processed and identifying the relevance of each category of data, even if 'stale'. This would allow maintaining a clear explanation of the relevance of the data collected. In addition, it seems relevant to consider the implementation of rules regarding the de-identification measures for certain data after a reasonable and relevant period of time. The idea of a specific data retention period specifically defined for Big Data Analytics may be considered an interesting approach to processing of data, provided that it is coupled with adequately defined processing rules and security measures.

In practice for principle 2: Data Quality
The entity collects sensitive information only if the individual provides explicit consent. Explicit consent requires that the individual affirmatively agree, through some action, to the use or disclosure of the sensitive information. Explicit consent is obtained directly from the individual and documented, for example, by requiring the individual to check a box or sign a form. This is sometimes referred to as 'opt in'.

[19] When data from other sources is used/embedded, the risk is that the used/embedded data is 'stale' (not continuously synchronized).

OECD privacy principle 3: Purpose Specification

The purposes for which personal data are collected should be specified no later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such other purposes which are not incompatible with the initial purposes

Big Data Analytics privacy & data protection challenge
In many Big Data Analytics cases, the purpose of (personal) data collection and processing is not clearly defined at the time of collection (or re-use) of the data. It is often the case that data is initially collected for a defined purpose, and that (re)use is contemplated at a later stage without necessarily having identified a rationale for this processing. Data is then retained after the initial purpose has been fulfilled, 'just in case'. This situation is becoming increasing common in Big Data Analytics: data controllers are more and more inclined to retain data they have collected without a new or further purpose.

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objective is applicable for this challenge.

- [GAPP 2.0] The entity provides notice about its privacy policies and procedures and identifies the purposes for which personal information is collected, used, retained, and disclosed.

Measure to address the privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 2.2.3] The entity's privacy notice is conspicuous and uses clear language.

In order to efficiently enforce this principle, it is essential that Big Data Analytics is immediately considered by data controllers when they collect personal data. Big Data Analytics cannot anymore be an 'afterthought' of the controller. The need for Big Data Analytics, whether for marketing purposes, etc. must be factored in and considered from scratch, resulting in including Big Data Analytics as a mode of processing personal data collected as part of the information provided to data subjects when collecting their information.

In practice for principle 3: Purpose Specification
The privacy notice is:

- in plain and simple language.
- appropriately labeled, easy to see, and not in unusually small print.
- linked to or displayed on the Website at points of data collection.
- available in the national languages used on the site or in languages required by law.

OECD privacy principle 4: Use Limitation

Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 of the OECD recommendation[20] (is related to the OECD privacy principle 3 - Purpose Specification) except:

a) with the consent of the data subject; or
b) by the authority of law.

Big Data Analytics privacy & data protection challenge
On the basis of OECD principle 3, data is generally collected for a defined purpose, e.g. marketing. Big Data operators propose the implementation of new processes which are unrelated to the initial purpose without having informed the data subject. Some business models for Big Data Analytics are based on sharing and selling (personal) data relating to data subjects who are not aware and have not consented to the further use of their data. In such cases, it is often argued that data shared and sold does not qualify as personal because it is 'de-identified'. However, using this data, combined with other data and via Big Data Analytics, an individual can be simply identified (even when the personal data is anonymized).

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objectives are applicable for this challenge.

- [GAPP 5.0] The entity limits the use of personal information to the purposes identified in the notice and for which the individual has provided implicit or explicit consent. The entity retains personal information for only as long as necessary to fulfill the stated purposes or as required by law or regulations and thereafter appropriately disposes of such information.
- [GAPP 7.0] The entity discloses personal information to third parties only for the purposes identified in the notice and with the implicit or explicit consent of the individual.

Measure to address the privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 7.2.2] Personal information is disclosed only to third parties who have agreements with the entity to protect personal information in a manner consistent with the relevant aspects of the entity's privacy policies or other specific instructions or requirements. The entity has procedures in place to evaluate that the third parties have effective controls to meet the terms of the agreement, instructions, or requirements.

Atos' Vision
As mentioned above in our analysis of Principle 3, including Big Data Analytics as a modus operandi of processing of personal data by including it at the very start of every project, by clearly advising data subjects of the possibility of further use of the data for a defined objective (or several defined objectives) at the time of collection seems to be the way forward to manage this issue.

[20] Organization for Economic Co-operation and Development (11 July 2013 - C(80)58/FINAL, as amended on 11 July 2013 by C(2013)79). Recommendation of the Council concerning Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data. Paris, France.

In practice for principle 4: Use Limitation
When providing personal information to third parties, the entity enters into contracts that require a level of protection of personal information equivalent to that of the entity's. In doing so, the entity:

- limits the third party's use of personal information to purposes necessary to fulfill the contract.
- communicates the individual's preferences to the third party.
- refers any requests for access or complaints about the personal information transferred by the entity to a designated privacy executive, such as a corporate privacy officer.
- specifies how and when third parties are to dispose of or return any personal information provided by the entity.
- The entity evaluates compliance with such contract using one or more of the following approaches to obtain an increasing level of assurance depending on its risk assessment:
- The third party responds to a questionnaire about their practices.
- The third party self-certifies that its practices meet the entity's requirements based on internal audit reports or other procedures.
- The entity performs an onsite evaluation of the third party.
- The entity receives an audit or similar report provided by an independent auditor.

OECD privacy principle 5: Security Safeguards

Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data.

Big Data Analytics privacy & data protection challenge
Some Big Data Analytic solutions put (personal) data in 'low cost' Cloud solutions. Low cost Cloud providers do not provide the same security safeguards (e.g. access to and protection of data) as the more mature Cloud providers or as your organization does. According to Gartner analyst Merv Adrian , even popular Big Data Analytics solutions like 'Hadoop' are definitely not secure; "at every layer of the stack, vulnerabilities exist, and at the level of the data itself there are numerous concerns".

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objective is applicable for this challenge.

- [GAPP 8.0] The entity protects personal information against unauthorized access (both physical and logical).

Measure to address the data privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 8.2.7] Tests of the effectiveness of the key administrative, technical, and physical safeguards protecting personal information are conducted at least annually.

Atos' Vision
Big Data Analytics policy should commit operators to implement security measures which are in line with the specificities of Big Data and the diverse nature of the data collected. In this respect, Big Data Analytics operators should consider the need or relevance of establishing pre-defined and / or innovative security measures on the basis of the data they are lead to process.

In practice for principle 5: Security Safeguards
Systems and procedures are in place to:

- regularly test the effectiveness of the key administrative, technical, and physical safeguards protecting personal informq ation.
- periodically undertake independent audits of security controls using either internal or external auditors.
- test card access systems and other physical security devices at least annually.
- document and test disaster recovery and contingency plan sat least annually to ensure their viability.
- periodically undertake threat and vulnerability testing, including security penetration and Web vulnerability and resilience.
- make appropriate modifications to security policies and procedures on a periodic basis, taking into consideration the results of tests performed and new and changing threats and vulnerabilities.
- periodically report the results of security testing to management.

21 Adrian, M. (21 January 2014). Security for Hadoop? Don't Look Now... Gartner Blog Network. http://blogs.gartner.com/merv-adrian/2014/01/21/security-for-hadoop-dont-look-now/.

OECD privacy principle 6: Openness

There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.

Big Data Analytics privacy & data protection challenge
Nowadays many organizations do not have a mature privacy & data protection management system in place. Therefore, developments and changes in processes involving personal data are not properly and promptly communicated to the individual, so the continuous validity of the initial consent is not guaranteed. Big Data analytics also raise the question of how data subjects can be effectively informed of the implementation of such Big Data processes when their data is being passed on and on to third parties.

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objective is applicable for this challenge.

- [GAPP 2.0] The entity provides notice about its privacy policies and procedures and identifies the purposes for which personal information is collected, used, retained, and disclosed.

Measure to address the data privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 2.2.1] Notice is provided to the individual about the entity's privacy policies and procedures (a) at or before the time personal information is collected, or as soon as practical thereafter, (b) at or before the entity changes its privacy policies and procedures, or as soon as practical thereafter, or (c) before personal information is used for new purposes not previously identified.

Atos' Vision
Defining a Big Data Analytics policy to define the fashion in which an organization intends to manage personal data in a Big Data Analytics context appears to be a way forward. Whether the organization in question defines a Big Data Analytics policy included in its own existing 'general' data protection / privacy policy or as a free standing policy should be up to the company. Defining a policy with regards to Big Data Analytics should include elements regarding the means implemented to process the data, the approach to data retention, the purpose with which the Big Data processing is implemented or the de-identification policy.

In practice for principle 6: OpennessThe privacy notice is:

- readily accessible and available when personal information is first collected from the individual.
- provided in a timely manner (that is, at or before the time personal information is collected, or as soon as practical thereafter) to enable individuals to decide whether or not to submit personal information to the entity.
- clearly dated to allow individuals to determine whether the notice has changed since the last time they read it or since the last time they submitted personal information to the entity.

In addition, the entity:

- tracks previous iterations of the entity's privacy policies and procedures.
- informs individuals of a change to a previously communicated privacy notice, for example, by posting the notification on the entity's Web site, by sending written notice via postal mail, or by sending an e-mail.
- documents that changes to privacy policies and procedures were communicated to individuals.

OECD privacy principle 7: Individual Participation

Individuals should have the right:

a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to them;
b) to have communicated to them, data relating to them
   i. within a reasonable time;
   ii. at a charge, if any, that is not excessive;
   iii. in a reasonable manner; and
   iv. in a form that is readily intelligible to them;
c) to be given reasons if a request made under subparagraphs 'a' and 'b' is denied, and to be able to challenge such denial; and
d) to challenge data relating to them and, if the challenge is successful to have the data erased, rectified, completed or amended.

Big Data Analytics privacy & data protection challenge
This principle addresses the right of an individual to be 'in control' of his/her personal data. Nowadays most organizations are not prepared for managing requests from individuals to provide the personal data collected by the organization and to enable the individuals to erase, rectify, complete or amend their personal data. Furthermore, most Big Data Analytics solutions are initially not built for direct interaction with end-users (and certainly not with individuals).

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objectives are applicable for this challenge.

- [GAPP 10.0] The entity monitors compliance with its privacy policies and procedures and has procedures to address privacy related inquiries, complaints and disputes.
- [GAPP 6.0] The entity provides individuals with access to their personal information for review and update.

Measure to address the privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 10.1.1] Individuals are informed about how to contact the entity with inquiries, complaints and disputes.

Atos' Vision
The challenge is difficult to address as Big Data Analytics creates a distance between the operators and the data subjects. Indeed, the data processed in the context of Big Data Analytics will generally not be directly collected by the operator and therefore, exercising one's rights in this context may be extremely complex unless the original collector retains a complete and accurate list of the entities with which it shares data – if at all possible. This issue also touches directly on the issue of the 'right to be forgotten' as data subjects will not be able to fully exercise this right if the original collector of the data does not have a clear vision of the entities which will have received the data it shared.

It is essential that Big Data Analytics operators offer the possibility for data subjects to efficiently and easily exercise their rights. In this respect, creating a platform or central repository of information where the data subject would be entitled to exercise their rights with a high level of granularity would permit effectively implementing the requirements.

In practice for principle 7: Individual Participation
The entity's privacy notice:

- describes how individuals can contact the entity with complaints (for example, via an e-mail link to the entity's Web site or a telephone number).
- provides relevant contact information to which the individual can direct complaints (for example, name, telephone number, mailing address, and e-mail address of the individual or office responsible for handling complaints).

OECD privacy principle 8: Accountability

A data controller should be accountable for complying with measures which give effect to the seven other OECD principles.

Big Data Analytics privacy & data protection challenge
This principle addresses the fact that organizations are accountable for securing the privacy and data protection (principles) during the entire lifecycle of the personal data processing. Nowadays most organizations do not have a sound privacy & data protection (risk) management system in place (e.g. based on the ISO/IEC 29100 standard) to design, implement and continuously improve the protection of privacy and data (see textbox Privacy by Design[22]).

Privacy & data protection objective for the challenge
The following GAPP privacy & data protection objective is applicable for this challenge.

- [GAPP 1.0] The entity defines, documents, communicates, and assigns accountability for its privacy policies and procedures.

Measure to address the data privacy & data protection challenge
The following GAPP measure is discussed to address the challenge.

- [GAPP 1.2.11] For each jurisdiction in which the entity operates, the effect on privacy requirements from changes in the following factors is identified and addressed:
  - Legal and regulatory;
  - Contracts, including service-level agreements;
  - Industry requirements;
  - Business operations and processes;
  - People, roles, and responsibilities;
  - Technology;
  - Privacy policies and procedures are updated to reflect.

Atos' Vision
In order to effectively implement the accountability principle, it is essential that the entity implementing Big Data Analytics processes has clear understanding of the purposes and means through which data is being processed in this context. The recommendations set out with regards to principles 1 to 7 above, if fully implemented, should allow for data controllers to be in a position to comply with the accountability principle. In addition to this, data controllers must be put in a position to obtain the relevant information regarding the fashion in which Big Data Analytics processes are implemented from their service providers and processors. Atos has initiated a road towards providing its customers more transparent information if they request it with regards to the processing of the Big Data Analytics processes they implement on their behalf.

In practice for principle 8: Accountability
The entity has an ongoing process in place to monitor, assess, and address the effect on privacy requirements from changes in the following:

- Legal and regulatory environments;
- Industry requirements (such as those for the Direct Marketing Association);
- Contracts, including service-level agreements with third parties (changes that alter the privacy and security related clauses in contracts are reviewed and approved by the privacy officer or legal counsel before they are executed);
- Business operations and processes;
- People assigned responsibility for privacy and security matters;
- Technology (prior to implementation).

Seven principles of Privacy By Design
1. Proactive not Reactive; Preventative not Remedial
2. Privacy as the Default Setting
3. Privacy Embedded into Design
4. Full Functionality
   => Positive-Sum, not Zero-Sum
5. End-to-End Security
   => Full Lifecycle Protection
6. Visibility and Transparency
   => Keep it Open
7. Respect for User Privacy
   => Keep it User-Centric

[22] Cavoukian, A. (August 2009 – revised January 2011). Privacy by Design, The 7 Foundational Principles. Ontario, Canada. Information and Privacy Commissioner of Ontario.

NIST and Big Data

In chapter 4 and 5 we addressed both the European data protection requirements and the 'soft law' requirements (e.g. defined in ISO/IEC 29100 standard and the GAPP standard). To derive the data protection requirements into Big Data Analytic architecture and processes a common Big Data (reference architecture) Framework is needed. Recently the National Institute of Standards and Technology (NIST) published the draft version of the 'NIST Big Data Interoperability Framework'. The specific privacy & data protection elements in the Framework and the presented privacy & data protection challenges (in chapter 5) are discussed in this chapter.

Big Data Interoperability Framework

On June 19, 2013, the NIST launched the NIST Big Data Public Working Group[23] (NBD-PWG). The goal of the Working Group is to develop a common Big Data framework (also addressed as the 'NIST Big Data Interoperability Framework'). The NIST Big Data Interoperability Framework consists of seven documents (called 'volumes') and addresses key Big Data topics drafted by five NBD-PWG subgroups[24]. The seven volumes are:

- Draft SP 1500-1 - Volume 1: Definitions
- Draft SP 1500-2 - Volume 2: Taxonomies
- Draft SP 1500-3 - Volume 3: Use Case & Requirements
- Draft SP 1500-4 - Volume 4: Security and Privacy
- Draft SP 1500-5 - Volume 5: Architectures White Paper Survey
- Draft SP 1500-6 - Volume 6: Reference Architecture
- Draft SP 1500-7 - Volume 7: Standards Roadmap

The NIST Big Data volumes are in mainly focused on United States organizations and are in an early draft. We believe that the NIST approach needs to be cross matched with the European Data Protection approach which may have implications on the way privacy must be managed.

Security & Privacy Subgroup

On April 6, 2015, the NIST Big Data Security & Privacy Subgroup published the draft version of Volume 6 regarding Security and Privacy. The scope of this volume includes among others the following topics:

- Security and Privacy Taxonomies;
- Security and Privacy Fabric of the NIST Big Data Reference Architecture (NBDRA).

The NIST Security and Privacy Taxonomies
The NIST Big Data Security & Privacy Subgroup has provided two taxonomies for Big Data security and privacy; the Conceptual Security and Privacy Taxonomy and the operational security and privacy taxonomy.

Conceptual Security and Privacy Taxonomy
The Conceptual Security and Privacy Taxonomy (see figure 1) contains four main Big Data security & privacy principles (called 'topics'):

- Data confidentiality topic: safeguarding
- the confidentiality of personal data.
- Data provenance topic: safeguarding the integrity and validation of personal data.
- System health topic: safeguarding the availability personal data.
- Public policy, social, and cross-organizational topics: safeguarding the specific Big Data and privacy & data protection requirements.

The first three topics are related to the traditional security principles; confidentiality, integrity, and availability (CIA). The fourth is specifically privacy & data protection oriented and addresses among others the social aspects of commerce, intellectual property, trans-border data flows and legal considerations.

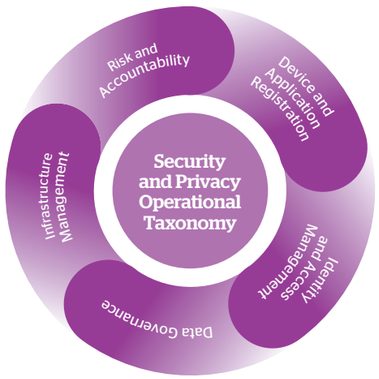
Figure 1 – NIST Conceptual Security and Privacy Taxonomy [NIST]

[23] For more information on the NIST Big Data Public Working Group, see: http://bigdatawg.nist.gov/home.php.
[24] The NBD-PWG subgroups are: Big Data Definitions & Taxonomies, Big Data Use Case & Requirements, Big Data Security & Privacy, Big Data Reference Architecture and Big Data Technology Roadmap.
[25] The concept of a "fabric" for security and privacy has precedent in the hardware world, where the notion of a fabric of interconnected nodes in a distributed computing environment was introduced. Computing fabrics were invoked as part of cloud and grid computing, as well as for commercial offerings from both hardware and software manufacturers (DRAFT NIST Big Data Interoperability Framework, Volume 4, Security and Privacy, Draft Version 1 April 6, 2015).

Operational Security and Privacy Taxonomy'
The 'Operational Security and Privacy Taxonomy' (see figure 2) contains practical checklists for using security & privacy methodologies within Big Data systems. The Operational Taxonomy addresses the following security & privacy topics for Big Data[26] :

- Device and Application Registration;
- Identity and Access Management;
- Data Governance: refers to administering, or formalizing, discipline (e.g., behavior patterns) around the management of data.
- Infrastructure Management;
- Risk and Accountability.

Figure 2 – NIST operational security and privacy taxonomy [NIST]



The NIST Big Data Reference Architecture
The NIST Big Data Reference Architecture Subgroup provided the NIST Big Data Reference Architecture[27] (NBDRA). "The NBDRA is a high-level conceptual model to facilitate an open discussion of the requirements, design structures, and operations inherent in Big Data. The NBDRA does not describe the system architecture of a specific Big Data system, but rather is a tool for describing, discussing, and developing system-specific architectures using a common framework of reference. The model is not tied to any specific vendor products, services, or reference implementation, nor does it define prescriptive solutions that inhibit innovation" [NIST].

The NBDRA contains five functional architecture components and two interwoven 'fabrics'.
The five functional architecture components[28] are related to the (technical) roles that exist in every Big Data system. The components are:

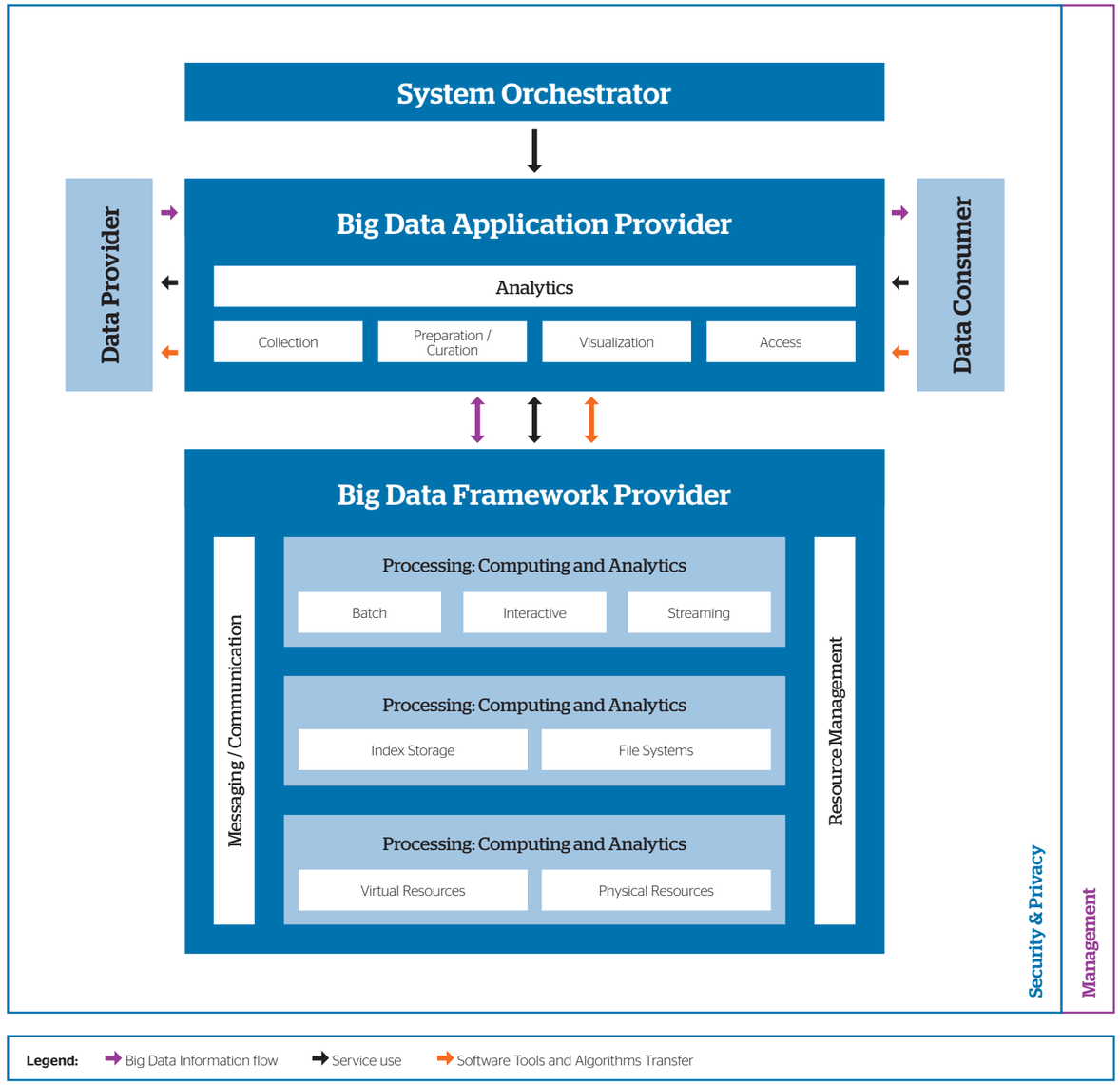
- The System Orchestrator: defines and integrates the required data application activities into an operational vertical system and provides the overall requirements that the system must fulfill, including policy, governance, architecture, resources, and business requirements and monitoring & auditing.
- The Data Provider: Introduces new data or information feeds into the Big Data system and makes data available to itself or to others.
- The Big Data Application Provider: executes the manipulations of the data lifecycle to meet the requirements established by the System Orchestrator.
- The Big Data Framework Provider: has the general resources or services to be used by the Big Data Application Provider in the creation of the specific application.
- The Data Consumer: Includes end users or other systems who use the results of the Big Data Application Provider.

The two NBDRA fabrics[29] provide services and functionality to the five functional components. The two fabrics are:

- Security and privacy: security and privacy issues affect all other components of the NBDRA. The security and privacy Fabric interacts with the System Orchestrator for policy, requirements, and auditing and also with both the Big Data Application Provider and the Big Data Framework Provider for development, deployment, and operation.
- Management: The management fabric encompasses two general groups of activities: system management and Big Data lifecycle management. System management includes activities such as provisioning, configuration, package management, software management, backup management, capability management, resources management, and performance management. Big Data lifecycle management involves activities surrounding the data lifecycle of collection, preparation/ curation, analytics, visualization, and access.

[26] For definitions and explanations of the terms see NIST Big Data Interoperability Framework, Volume 1, Definitions and Volume 4, Security Draft Version 1, April 2015 and Privacy, Draft Version 1, April 6, 2015.
[27] Described in NIST Big Data Interoperability Framework, Volume 6: Reference Architecture, Draft Version 1 April 6, 2015.
[28] The definitions of the NBDRA components are provided by NIST (see also reference 14 and 15).
[29] See reference 16

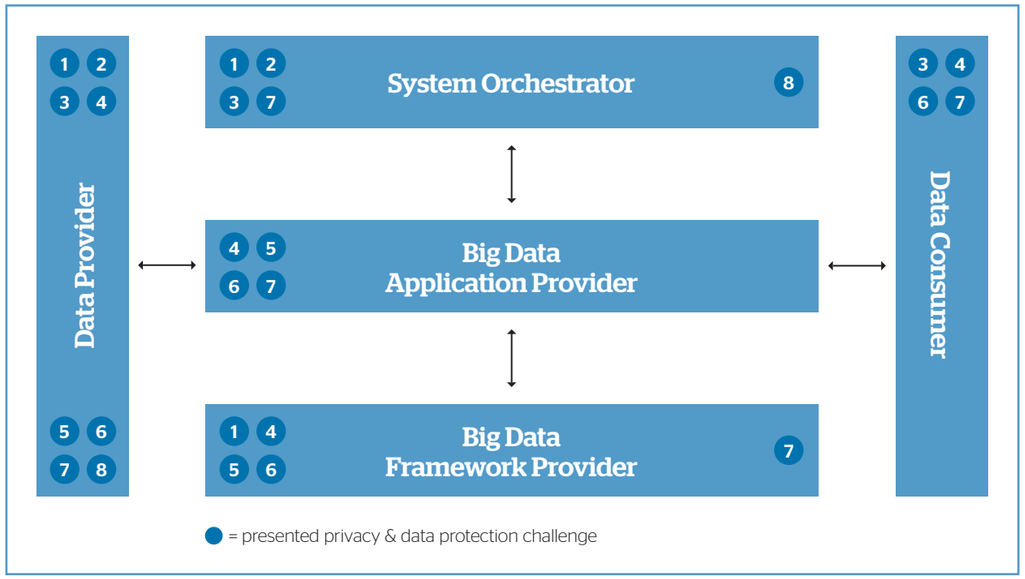
Figure 3: NIST Big Data Reference Architecture [NIST]



Mapping challenges to NDBRA

To provide an initial aid for implementing relevant (technical) privacy & data protection solutions[30] (see table 2) in Big Data systems we have mapped the privacy & data protection challenges (see chapter 5) in the simplified NIST Big Data Reference Architecture (NBDRA) model. The mapping is presented in figure 4.

Figure 4: Mapping privacy & data protecting challenges against NBDRA



[30] Described in DRAFT NIST Big Data Interoperability Framework, Volume 4, Security and Privacy, Draft Version 1 April 6, 2015).

Table 2 – NBDRA Roles mapped against Operational Security and Privacy Taxonomy [NIST]

NBDRA (Technical) Roles **Operational Taxonomy of security and Privacy Topics**	**System Orchestrator**	**Data Provider**	**Data Consumer**	**Application Provider**	**Framework Provider**
Device and Application registration					
Device, User, Asset, Services, Applications Registration		X			
Security Metadata Model	X				
Policy Enforcement	X				
Identity and Access Management					
Virtualization Layer Identity					X
Application Layer Identity		X	X	X	
End User Layer Identity Management		X	X		
Identity Provider					X
Data Governance					
Encryption and Key Management					X
Isolation / Containerization					X
Storage Security					X
Data Loss Prevention, Detection	X				
Web Services Gateway			X	X	
Data Transformation				X	
Data Lifecycle Management	X				
End Point Input Validation		X			
Digital Rights Management		X	X	X	
Trust	X	X	X	X	X
Openness	X				
Fairness and information ethics	X				
Infrastructure Management					
Threat and Vulnerability Management	X				
Monitoring, Alerting	X	X	X	X	X
Mitigation	X				
Configuration Management	X				
Logging					X
Malware Surveillance and Remediation	X				
Network Boundary Control					X
Resiliency, Redundancy and Recovery	X				
Risk and Accountability					
Accountability	X				
Compliance	X				
Forensics	X				
Business Risk Model	X				

Conclusion

With Big Data Analytics comes Big Responsibility. That Big Data Analytics is emerging is evident, not only for organizations, but also for politicians, regulators and consumers. Therefore the potential for Big Data Analytics is very high, for organizations and consumers alike. The new EU Regulation is a great step forwards towards supporting business innovation by using Big Data Analytics solutions.

However, it is our conviction that considering privacy & data protection principles at the very start of the implementation of Big Data Analytics systems and projects can allow organizations to not only fully unleash the capabilities of Big Data Analytics but can allow them to differentiate themselves from competitors who may be exposed to complaints and potential sanctions.

Therefore, to significantly improve business efficiency and innovation, organizations must act within the spirit of the 'key privacy & data protection principles' and must offer privacy and data protection compliant to the Regulation and user friendly Big Data Analytics solutions by implementing applicable controls.
The design and implementation of privacy & data protection controls is a continuous process. Therefore, organizations must establish a privacy & data protection (risk) management system (e.g. based on the ISO/IEC 29100 standard) to identify, mitigate and manage related risks in line with the 'risk appetite' of the organization.

In order to achieve this objective, it will be necessary to consider and implement new, innovative and efficient solutions to guarantee a level of protection to personal data processed in the context of Big Data Analytics. Solutions have to combine both legal instruments and technological tools. This White Paper describes trends and makes proposals towards the development and use of such tools, mainly on the basis of legal principles. Technological solutions are being developed every day and Atos intends to focus even more on these potential solutions in an even more technology-oriented White Paper in the months to come.

"To significantly improve business efficiency and innovation, organizations must act within the spirit of the key privacy & data protection principles"

Colophon

Acknowledgements

The author would like to thank his Atos Consulting Benelux & The Nordics colleagues (in alphabetical order): Rutger Fichtinger, Alistair van Heezik, Till Kolloge, Angeline Lepine-Luiten, Johan Pater and Robin Zondag, for their support and contribution in drafting the white-paper. Furthermore I would like to thank the members of this Atos white Paper review group for reviewing the white Paper and providing constructive feedback (in alphabetical order):

- Olivier Maas (Atos Worldline and Member of the Atos Scientific Community)
- Paul Oor (Atos Chief Security Officer Benelux & The Nordics)
- Paul Rakke (Atos Global Big Data Community)

About the authors

Henk Brandon
Henk is Managing and Principal Consultant at Atos Consulting Benelux & The Nordics.

Atos Consulting helps you realize your future-proof efficiency, agility and improved top line. Our key to success is coupling the right strategy, process design and innovation with IT. We are motivated to work with you and your staff to achieve this challenge. Our clients view us as leaders in commitment to implementation.

Henk works within Atos Consulting in the Governance, Risk and Compliance competence group. From this competence group Henk works at the interface between business and IT and is optimizing and securing the organization, people, processes and ICT. He assists organizations, both public and private institutions, in managing the operational business and IT risks.

Lionel de Souza
Lionel is Atos Group Chief Data Protection Officer and Member of the Atos Scientific Community.

Lionel is tasked with developing, implementing, coordinating and enforcing, at group level, the privacy and data protection principles which Atos, in compliance with applicable laws, has enshrined into its culture.

For that purpose, Lionel works with a coordinated team of data protection experts, (data protection officers and data protection legal experts) to assist all stakeholders of the company to ensure they implement their processes in compliance with applicable data protection legislation and provide customers with adequate, innovative and easy to handle solutions.

The Atos Scientific Community is a network of some 100 top scientists, representing a mix of all skills and backgrounds, and coming from all geographies where Atos operates. Publicly launched by Thierry Breton, Chairman and CEO of Atos, the establishment of this community highlights the importance of innovation in the dynamic IT services market and the need for a proactive approach to identify and anticipate game changing technologies.

About Atos

Atos SE (Societas Europaea) is a leader in digital services with 2014 pro forma annual revenue of €10 billion and 86,000 employees in 66 countries. Serving a global client base, the Group provides Consulting & Systems Integration services, Managed Services & BPO, Cloud operations, Big Data & Security solutions, as well as transactional services through Worldline, the European leader in the payments and transactional services industry. With its deep technology expertise and industry knowledge, the Group works with clients across different business sectors: Defence, Financial Services, Health, Manufacturing, Media & Utilities, Public Sector, Retail, Telecommunications and Transportation.

Atos is focused on business technology that powers progress and helps organizations to create their firm of the future. The Group is the Worldwide Information Technology Partner for the Olympic & Paralympic Games and is listed on the Euronext Paris market. Atos operates under the brands Atos, Atos Consulting, Atos Worldgrid, Bull, Canopy, and Worldline.

For more information:
Please contact **henk.brandon@atos.net** or **lionel.desouza@atos.net**

atos.net