International Journal of Core Engineering & Management
Volume-6, Issue-06, 2019 ISSN No: 2348-9510

STARTING WITH THE DATA PURPOSE FOR IDENTIFYING SECURITY MEASURES

Anand Athavale
Independent Researcher
Decades of Data Management experience
San Jose, United States of America

Abstract

This article explores the fundamental aspect of relying on purpose of data to define processes and the elements, to effectively assess, improve and monitor data security measures. Effective Data security posture helps reduce probability and impact of a cyber or a compliance breach. This article takes a methodical approach for first defining the purpose of the data, whether already created, or, about to be created and then introduces the process of identifying elements of data security posture from the purpose definition. Data purpose definition and identification is the stepping stone for effective security management. A clearly defined data purpose, automatically starts revealing the necessary elements to measure, improve and monitor security of the data. After reading this article, any information, compliance, and security office team member will be equipped to map the data purpose communicated by the data owners and operators to them to their objective of securing the data.

Index Terms – Information Security, Data Compliance, data security measures, Ransomware resilience, data purpose

I. INTRODUCTION

As with many things, all data is not created equal. Apart from many other aspects and attributes that separate one data item from the other, there is a fundamental defining aspect, which then directs the process for security management. Like many things in life, data has a purpose, or at last needs to have one. It can then help define the measures needed for its security. Purpose of each data at a generic level is providing information and keeping records. However, more specifically, it is different in terms of many aspects such as area it falls into such as finance, weather, geological and more. But the specific purpose leads to defining certain traits or attributes, which then help identify the process for the security measures. To illustrate further, think of a project you may be starting in your organization. The project starts with defining an objective. That objective further determines the stake holders and contributors. That project may be one time, such as moving your data center, or, on-going like creating continuous performance measurement system. Like any project objective, data purpose helps determine certain elements which contribute to defining security measures. Simply defining the purpose itself uncovers many aspects as we will learn in this article.

II. DATA PURPOSE AND DATA PURPOSE STATEMENT

The purpose of data can become very complex to define accurately in detail. But even before we consider purpose of data, we need to acknowledge that, what gets referred to as Data or information also can be often left to interpretation. Is anything which gets stored or communicated qualifies as data? The answer is "it depends". For example, would you consider the operating system binaries on various servers as "data"? Most likely not for the purpose of this discussion. Even if one did so, the challenges for data security measures are more applicable to the data that is produced using various applications. Operating systems binaries themselves are the backbone of such applications. We may call those infrastructure or supporting systems for data. However, in certain programming languages, data can very well be part of the code, and hence could qualify as data. But those are not the target of the discussion. This clarifies the context for this article.

A basic statement of purpose helps us discuss and define the data security posture deciding elements. The idea here is whether it is a single data item, or a collection, it is created by a set of authors, or producers, and it is meant for a set of readers or consumers. In other words, the producers and consumers for any data item, or collection of data items, is less generic and more specific.

Let's take a look at a simple statement of purpose for this very article.

This article is meant to illustrate the process for identifying data security measures for most of the data, to the technical and operational community grappling to define, measure and maintain security of data, to combat increasing threat of cyber-attack and to address and fulfil related regulations.

III. THE "What it is"

First part identifies the data item itself in a manner. The words, "This article", tells what it is. So, content wise, it is an article. For the technology or operational side, it tells that it is some type of a document. I also know that I am writing it in a word application, so it is an office application document format [3]. In storage sense, this document is a "file." Hence, it most likely is going to be under a "folder" or in some collection like a folder.

So far, we have concentrated on the highlighted section of the purpose statement and we identified some portions of data security posture elements looking at the "What it is" part the purpose.

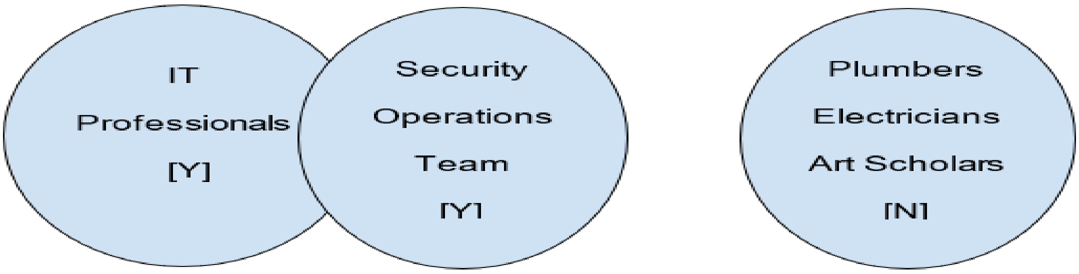
"This article is meant to illustrate the process for identifying data security measures for most of the data, to the technical and operational community grappling to define, measure and maintain security of data, to combat increasing threat of cyber-attack and to address and fulfill related regulations."

IV. THE "What it covers"

The "objective" of the article is defined as "illustrating the process for identifying data security measures." It narrows down the "topic" of this data. If there were keywords to associate with this data item, most likely than not, "data security measures" would come out to be a winner. Once a topic is identified, it further crystalizes the "Who" part. This happens because specific set of people would be related to, or, interested in certain topics.

V. THE "Who it is meant for"

The purpose statement brings put the important aspect of the audience or the consumer of the data item, which is crucial from security measures point of view. For this very article, it is a the "technical and operational community grappling to define, measure and maintain security of data." While the intended consumer or the audience here is rather broad, it is clearly "defined". If you consider a set theory, it also automatically defines who it is not for. It is not for medical or arts scholars as an example. It is probably not that helpful to electrical or water systems professionals. This may be obvious, but this very thought process will help us when we define data security posture elements of exposure control. Exposure control is a key tool or means of securing the data [1].



It is important to note that sometimes, the audience may not be clearly defined. So, in the absence of that, the focus would automatically be on the other elements of the data purpose statement, to figure out the audience and put the necessary exposure control. This could also be referred to as the "Who" of the data, however, it can have more than one meaning. In certain scenarios, the producers of data matter equally, or, more than the consumer. The "Who" part can also refer to the owner of the data in case of personal data as illustrated in general data protection regulation [2]. In the modern cloud services world, the association of the data to the producer and consumer becomes extremely difficult to determine because the operations sometimes can be carried out by services and service principles. Service and service principles do not always have a direct association to single person or even a group of persons, compared to active directory or similar identity systems.

While data purpose is supposed to be defined first, data creation does not always wait for the security measures to be fully formed. There are many elements beyond those covered by the data purpose, like the attributes of the producers and data location types. These attributes relate not only to setting security measures, but also create requirement of constant monitoring and improvement. In this example, in the case that the portion "to the technical and operational community" was not mentioned, here is how it would have been determined. The "what" part about data security posture process illustration, would have automatically reduced the audience to only those who deal with data and data security in some way or form. The portion "grappling to define, measure and maintain the data security posture" portion would have further narrowed it down to security and IT practitioners. Thus, data purpose still plays a vital role in defining the process for data security posture.

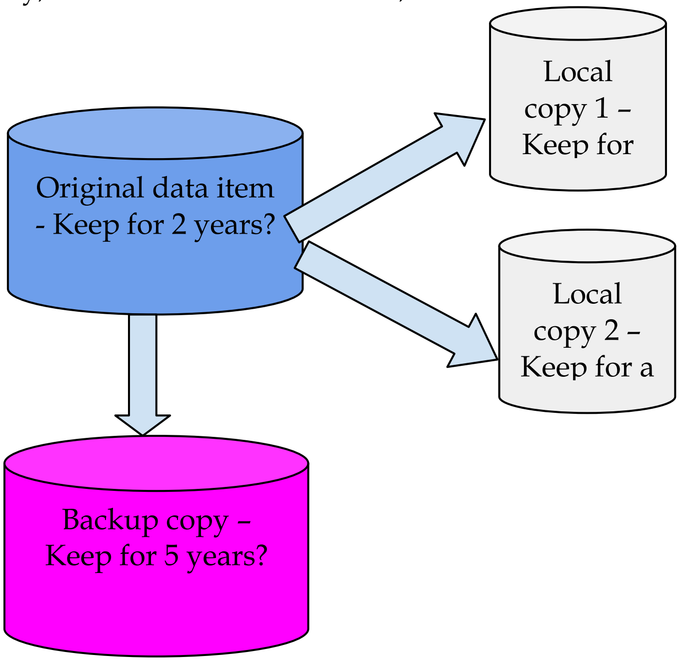
VI. THE "Why"

Lastly, the purpose statement defines the why of the document. The why helps identify the criticality and importance of the document. It also touches a little bit on how long this document needs to be stored or retained. One would say, at the source, where it is published, it needs to be preserved for a few years give or take. But, if anyone downloads it and uses it, then that copy can be deleted once the purpose is served. So, maybe after a month or so, one may not find use for the downloaded copy. But, if there is an updated "version" of the paper, then, it will again need to be downloaded and kept for maybe another month or so. In the "why" of the purpose, we have uncovered quite a few aspects of the security measures.

One was the criticality, which would define the intensity and thoroughness needed for securing this data item. It is important to acknowledge that while looking at the "why" of the data purpose, many other factors including the actual content of the data also weighs in on the level of effectiveness and a number of measures required for securing any data item.

Second was the retention time for this item, depending on where it was stored. Thus, in the same contemplation about the why, we uncovered that same data items may exist in more than one location. As such, data location is one of elements that too influences the data security measures required. All these elements deserve consideration in a dedicated fashion and are too vast to cover under this article.

Third, that data items move and replicate, as a "copy". Other than just for having it local for immediate access, there may be more reasons for creating a copy. For example, a copy may be required for backup purposes, in case of corruption, unintended modifications or loss and deletion [4]. Another reason for a copy may be to meet certain regulations for specific types of content to ensure it is retained for a certain period, but with less immediate access [5]. In the old days, such media used to be tape, or, CD-ROMs. In the modern world, there are what is referred to as Write Once Read Many, or WORM media locations, which can include cloud storage used for archive.


Original data item - Keep for 2 years?
Local copy 1 – Keep for
Local copy 2 – Keep for a
Backup copy – Keep for 5 years?

Last one was about fluidity and changes to the data item, introducing the technical concept of a "version". Version is basically another incarnation of the same data item with some modifications but tracked meticulously usually with some kind of labelling or numbering system [6]. Version tracking sometimes is built into the storage systems, while it could also be maintained manually.

Irrespective of whether the versioning is automated, or manual, it still affects the data security measure identification process. Again, version control itself is a vast topic to cover in this article. But it does show that as basic step as defining a data purpose is a useful first step in identifying related security measures needed.

VII. CONCLUSION

In this article of data security measure identification, we identified the benefits of defining a data purpose. A simple data purpose statement itself can give a good head start on figuring out data security measures required for any data item. Of course, identifying security measure is complex, since there are many more unwritten elements from the data purpose aspect. However, defining data purpose gives us opportunity to ask ourselves more questions which help uncover more elements such as exposure control, classification, and labelling and so on. Directly jumping to set of measures for securing data may become more iterative than needed, and often may leave out important stake holders in the process. Today securing data seems to be a responsibility pushed to very small number of roles and specific teams within an organization. But those isolated teams may not necessarily know all the required aspects of data to then employ the right and sufficient security measures.

Just to give a glimpse as an example, data location might focus more on the storage aspect. While security gets discussed for storage layers, from implementation aspects, it will be more specific to storage administrators and less to data producers and security operations. However, monitoring and improvement of security measures falls more on the shoulders of security operations commonly known as InfoSec team.

REFERENCES

1. Ji-Won Byun, Purpose Based Access Control of Complex Data for Privacy Protection, CERIAS and Department of Computer Sciences, Purdue University, (2005) https://www.cs.purdue.edu/homes/ninghui/papers/purpose_sacmat05.pdf (June, 2019)
2. Rick Krutzen, Why and how we should care about the General Data Protection Regulation, Taylor and Francis Online (2019), https://www.tandfonline.com/doi/full/10.1080/08870446.2019.1606222, (June, 2019)
3. Aaron Peters, Everything you need to know about file formats and their properties, Make Use Of (MUO) (2017), https://www.makeuseof.com/tag/file-format-properties-explained/ (May, 2019)
4. Eric Hibbard, Storage Security: Data Protection, Storage Networking Industry Association, (March 2018) https://www.snia.org/sites/default/files/security/SNIA-Data-Protection-TechWhitepaper.pdf (June, 2018)
5. Chuck Cook, Factors to consider when developing a data retention policy for your business, (Feb 18, 2015) https://www.renovodata.com/blog/2015/02/18/developing-data-retention-policy (July, 2019)
6. Lorna Maguire, Version Control Guidance, (July, 2017), https://www.abdn.ac.uk/staffnet/documents/policy-zone-information-policies/UoA_Version%20Control_July%202017.pdf (July, 2019)