Data Analytics Policy

|  |  |
| --- | --- |
| Policy Area | IT Policy Library |
| Approved Date | December 31, 20XX |
| Approved By | Policy Committee |
| Effective Date | January 1, 20XX |
| Current Version | 1.0 |

# I. Overview

Data analytics helps ABC Company increase revenues, improve operational efficiencies, optimize marketing campaigns, enhance customer service, respond more quickly to trends, gain a competitive edge in our marketplace, and enhance business performance.

# II. Purpose

The purpose of this policy is to provide a foundation for the development and implementation of necessary security controls to protect data according to its value and/or risk. Data that is analyzed may originate from a mix of internal systems and/or external sources. The data may consist of either historical records or new information that has been processed for use in real-time analytics.

# III. Scope

This policy applies to all Staff involved in the collection, preparation, use, storage, and disposal of analyzed data.

# IV. Policy

A. Background

Data analytics refers to techniques and processes used to extract and categorize data for the purpose of identifying and analyzing data and patterns. Data analytics involves the discovery, interpretation, and communication of meaningful patterns in data. Data analytics may include data mining, which can include sorting through large volumes of data to identify trends, patterns, and data relationships. Predictive analytics seeks to predict behavior and future events.

Software analytic applications, a type of business intelligence software, use historical data to provide information and tools that allow management to make informed business decisions. Self-service business intelligence tools allow users to create and run their own ad hoc queries and reports.

Data analytics typically involves the following phases:

* Data collection (data ingestion). Identify information needed and resources to collect or assemble the data.
* Data preparation. Perform data profiling and cleansing to ensure the data is in an appropriate format, type, length, etc. Duplicate records are eliminated and data governance policies are applied to ensure that the data adheres to required standards. Issues involving large numbers of data records are addressed.
* Data analytics. Predictive modeling and analytics software and programming languages such as Python, Scala, R and SQL may be used. On large data volumes, the software may initially run against a subset of the data to ensure the data and model are working as desired.
* Data communications. The results generated by analytical models are communicated to users and management.

A data lake is data stored in its natural or raw format. A data lake may include raw copies of data as well as transformed data used for reporting, visualization, analytics, machine learning, etc. Data lakes may include unstructured data (e.g. documents, e-mail messages, PDF files, etc.) in natural format, structured data from relational databases, as well as binary data such as images, audio, and video. Cloud storage services such as Microsoft Azure Data Lake and Amazon S3 or a distributed file system such as Apache Hadoop may be used.

Data virtualization allows an application to retrieve and manipulate data without knowing technical details (e.g. formatting, location) about the data.

Big data applications collect, manage, and process large volumes of data that are typically beyond the ability of commonly used software. Big data has the following characteristics:

* Volume – large amount of data.
* Variety – type and nature of the data
* Velocity – speed at which the data is generated and processed
* Veracity – data quality and value

B. Security Controls

The Risk Management Officer (RMO) shall ensure risks are properly identified and managed through a three-step process (for more information see the Risk Management Policy):

* Risk Assessment – Identify assets, threats to the assets, and vulnerabilities that exist as a result of the threats.
* Risk Analysis – For each risk, identify the likelihood and impact on our organization.
* Risk Treatment Plan – Develop controls and other safeguards designed to mitigate, eliminate, or transfer risks. It is important to treat risks in a cost effective and efficient manner.

The Chief Security Officer (CSO) shall ensure adequate safeguards are in place to ensure the security of Sensitive Information and Information Systems. Safeguards shall protect against electronic threats as well as unauthorized access, disclosure, copying, use or modification of systems and data. Data collected and processed shall be limited to the amount and type of information specifically required for an identified purpose. For more information on applicable security controls and requirements see:

* Access Control Policy
* Anti-malware Policy
* Application Implementation Policy
* Approved Application Policy
* Backup Policy
* Capacity and Utilization Policy
* Configuration Management Policy
* Data Classification Policy
* Data Privacy Policy
* Encryption Policy
* Firewall Policy
* Information Security Policy
* Logging Policy
* Network Security Policy
* Physical Access Policy
* Privacy Policy
* Retention Policy
* Security Controls Review Policy
* Security Policy
* Software Development Policy
* Third Party Service Providers Policy
* User Privilege Policy

# V. Enforcement

Any Staff member found to have violated this policy may be subject to disciplinary action, up to and including termination.

# VI. Distribution

This policy is to be distributed to all ABC Company Department Heads, the Risk Management Officer, and the Chief Security Officer.

**Policy History**

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Description | Approved By |
| 1.0 | 1/1/20XX | Initial policy release |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**References:**

COBIT EDM03.02, EDM03.07, APO10.05, APO12.02, APO13.07, APO14.10, BAI01.03, BAI04.05

GDPR Article 9, 10, 25, 32

HIPAA 164.308(a)(1)(ii)(A), 164.308(a)(1)(ii)(B), 164.308(a)(3)(ii)(B), 164.308(b)(4)

ISO 27001 6.1.3, 9.1, A.8.2.1, A.9.1.2, A.9.4.1, A.12.1.3, A.12.2.1, A.12.4.1, A.12.6.2, A.13.2

NIST SP 800-37 3.3, 3.4, 3.5

NIST SP 800-53 AC-1, AC-3, AC-24, PM-29, RA-2, SC-13, CP-2

NIST Cybersecurity Framework ID.RA-6, PR.AC-4, PR.DS-1, PR.DS-2, PR.DS-4, PR.DS-5

PCI 3.1, 7.1, 12.2