- **How to Ensure Data Integrity in Your Organization**

We are living in an era of data and analytics. It is the time when organizations are consuming, generating, modifying, and exchanging an insane amount of data.

Now, failing to keep track of data that is being consumed, generated, modified, exchanged, or deleted within your organization can have a major impact on your business decisions. That is why it is critical to preserve data integrity in your organization.

Especially if you are a data-driven organization, data integrity will act as a crucial asset that will prevent you from putting your business in jeopardy by making reckless business decisions.

That is why I have written this blog in which I will talk about:

- What Is Data Integrity?
- Why Is Data Integrity Important?
- Data Integrity Threats
- How to Know if Your Data Has Integrity
- 13 Ways to Ensure Data Integrity in Your Organization
- Ensuring Integrity for Open Data

What Is Data Integrity?

Data integrity is the process of maintaining and assuring the accuracy and consistency of data throughout the data lifecycle.

The idea is to ensure your data is reliable, accurate, stored in the best way possible, and does not change when you modify, transfer, or delete it.

For better understanding, we can split data integrity into two parts:

- **Physical Data Integrity:** deals with how we store and access data in our organization (using on-premise servers or the cloud databases you are remotely connected to). Ensuring the physical security of devices and having a disaster recovery plan are our major concerns in this part.
- **Logical Data Integrity:** is all about preventing human error to prevent the loss of data integrity.

Why Is Data Integrity Essential?

Organizations are heavily relying on data and analytics for making crucial business decisions these days. Now imagine how a business decision based on data that is unreliable and inaccurate will impact your organization.

Not only will you potentially lose millions, but your organization’s reputation will also face a setback.

According to a report from KPMG, 92 percent of senior executives have the same concern. They are worried about the negative impacts of data and analytics on their organization.

Also, the chances of your data getting compromised increase tenfold without data integrity, and compromised data is of no use for companies.

Consider the example of a company dealing in dry ice temperature data loggers. Imagine what will happen if their data gets compromised due to a data integrity breach. Shipments will get spoiled, and the company will lose millions. Sometimes it can even put someone's life at stake.

That is why paying attention to data integrity is essential. It is the basis for reliable insights. You can vouch for the searchability, traceability, and quality of data only if you are aware of where it has been generated, where it has been transferred, and how many times it has been modified.

In simple words, data integrity is always the best choice to maintain the quality of data throughout your organization. It can save you from a lot of trouble.

Common Reasons Behind the Loss of Data Integrity

1. Human Errors

Data is unavailable because it was accidentally deleted, inaccurate because of typos and wrong data entry, or incomplete because of the negligence of your employees. While sometimes these errors are unintentional, there are times when there is malicious intent behind them.

2. Transfer Errors

This could be a data transfer to a wrong address or compromise when the file was being transferred between two devices. As a result, data is no longer of any use to your organization.

3. Cyber Threats

Examples would be bugs, spam, malware, and other cyber threats that are targeted at your organization to breach your data integrity.

4. Security Issues

This includes security loopholes or misconfigurations that can be exploited by cyber terrorists to breach your system and compromise your data integrity.

5. Hardware or Infrastructure Issues

This can be outdated hardware, physical compromise to devices, or infrastructure in which little attention is paid towards security.

The issue is only a few of these data integrity threats can be prevented with the help of data security. Unless you have a strong strategy for maintaining data integrity at your workplace, it will become hard for you to overcome these issues and protect the integrity of data.

How to Know if Your Data Has Integrity

Following are some aspects from which you can determine the integrity of your data:

Availability: Your data is available whenever you need it, and wherever you need it.

Integrity: Your data was recorded as it was observed, and at the time it was executed.

Reliability: Your data is accurate, free from errors, and adheres to protocols. You can blindly trust it for making crucial business decisions.

Comprehensibility: Your data is easily accessible, comprehensible, and permanently recorded.

Originality: All the original entries of your data have been preserved, and you can access them when you need them.

Transparency: Your data clearly demonstrates when it was created and stored, who created and stored it, and what the data is all about.

14 Ways to Ensure Data Integrity in Your Organization

1. Data Entry Training

One of the biggest hurdles that businesses face in maintaining data integrity is that their employees don't even know how to preserve it. Hence, offering data entry training to them can be the best way to get started.

Start with training your employees on how to enter and maintain data and delegate them with the responsibility of preserving the Data Quality. It will make sure everyone in your team is putting effort towards maintaining data integrity.

2. Validating Input and Data

This is important, especially if your data is being supplied by an unknown source like an end-user, another application, or third-party sources. You need to verify and validate the data inputs to ensure they are accurate.

Apart from this, you also need to validate data from time to time so that you can be assured that data processes have not been corrupted.

3. Removing Duplicate Data

Duplicate data is one of the biggest reasons behind the breach of data integrity as it often causes ambiguity and leads to malicious errors. So, identifying and removing duplicate data on time is important.

While big companies have a dedicated staff for this purpose, you can use duplicate file cleaners if you are a small organization and do not have dedicated resources for this role. Here are some of my favorites:

- BleachBit
- FSlint
- rmlint
- Rdfind
- GDuplicateFinder

The best thing about the list above is that all these tools are open source, which means you can not only use them for free but also enjoy support from the community.

4. Backing Up Data

Even if you are taking necessary measures to preserve data integrity in your organization, you still cannot risk losing permanent data. Hence, I would advise you to regularly take a backup of your data.

You need to understand that in the case of cybersecurity attacks that are most likely to happen, data backup will play a critical part in ensuring we can bring everything back to normal and preserve the integrity of data.

Also, you must be careful while choosing a cloud storage and backup solution for your organization. Most providers are not what they claim to be, and a wrong choice can put you at great risk. So, choose wisely.

In case you are wondering, this comparison of the best cloud storage services can help you make the right decision.

5. Using Access Controls

Chances of a data integrity breach increase tenfold when there is no access control in your organization. It often leads individuals without any organization access and malicious intent to gain access to your data and do grievous harm to your organization.

That is why it is important to exercise access control in your organization. Implement a least privilege model and give access to only those users who need it so that there is high-level control and data integrity can be preserved.

In case you are wondering how to exercise access control in your organization, this guide to using access control in your organization can help.

6. Keeping an Audit Trail

Whenever there is a data breach, it is important to track its source. That is why keeping an audit trail is crucial to preserve data integrity.

An audit trail will provide the organization with breadcrumbs that will highlight the source of the problem so that you can resolve it on time.

Here are the qualities an ideal audit trail should have:

- The audit trail should be generated automatically
- Users should not be able to tamper with the audit trail
- Every event of the audit trail (creation, deletion, modification) should be tracked and recorded
- Every event should be aligned with users so that you know who accessed the audit trail
- Every event should be time stamped so that you know when the event took place

7. Establishing Collaboration in the Organization

Another way to maintain data integrity in your organization is to keep all the members of your organization on the same page. They should know who is creating, modifying, and transferring the data and when.

That is why you should ensure collaboration in your organization. All members should be able to collaborate and work together. Use emails, business phone systems, conference calling services, or tools like Microsoft Teams for collaboration and communication. In the end, everyone should be in the loop.

Not only will it allow you much-needed control over your data but also keep you on the same page with the rest of the members of your organization.

In case you are wondering, here are15 toolsto ensure collaboration in your organization.

8. Performing Penetration Testing and Security Audits

Another efficient measure to check data integrity is penetration testing, i.e., having an ethical hacker try and hack into your company's database and find its vulnerabilities. It will help you find where you are lacking and fix the problems on time.

You can also perform security audits as they only require the involvement of internal sources. While doing so, make sure your security audit covers the following aspects:

- Cybersecurity practices for employees
- Information that is critical for your organization
- Practices hackers can follow to breach your data
- Security measures to identify and prevent potential hacking activities

9. Actively Performing Volume and Stress Tests on the Database

Data integrity breaches also happen when the hardware is not able to cope with the information it has to process. These situations often lead to technical attacks and server malfunctions.

Bad code and poor configuration can also be another reason as attackers can use them to brute force login screens and acquire user passwords.

That is why it is important to perform volume and stress tests on the database from time to time so that all these issues can be tracked and fixed on time and data integrity remains preserved.

You can refer to this list of 40-plus database testing tools if you are wondering how to test your database for stress and volume.

10. Encrypting Your Data

Encryption can also be an effective measure to preserve data integrity in your organization. It ensures that even if someone is able to access your data, they cannot read it without the decryption key.

It works well in scenarios when attackers can easily acquire files stored in your database by means like stealing the server or downloading files by hacking into your database.

11. Enabling SSL Encryption on Your Websites

Websites are one of the most common reasons cyber criminals may get access to your data and breach your data integrity. Man-in-the-middle and downgrade attacks are very common on them.

That is why you should enable SSL encryption on your website. It will encrypt your communication and eliminate all these threats.

Here is how you caninstall an SSL certificate on your website.

12. Developing Process Maps for Critical Data

For preserving the data integrity, you also need to have control over how and where the data is being used and who is using it. It will keep a check on your organization's data and prevent its misuse.

That is why it is essential to develop process maps for critical data so that your organization has greater control over its use. It will also help you implement proper security measures and regulatory compliances.

Wondering how to develop a process map for your critical data? You can refer tothese 11 steps to process mapping.

13. Promoting a Culture of Integrity

Data integrity is not just about taking some precautionary measures but also about establishing a culture of transparency, honesty, and integrity.

Your team members need to be honest about their work and take ownership of their data. They also need to report instances in which other members break the rules and fail to fulfill their responsibilities.

These are the little steps that will make sure your entire organization stays on track, and the data integrity of your organization remains intact.

14. Paying Attention to Cybersecurity

Cybersecurity will play a critical role in preserving data integrity in your organization. It will ensure no one can access essential data without your permission, and chances of a data breach can be prevented.

So, ensuring cybersecurity in your organization should be your top priority. You must draw strict policies, use advanced security tools, and take necessary security measures.

Ensuring Integrity for Open Data

We learned how to ensure data integrity in your organization. But ensuring and maintaining integrity for open data is an entirely different ballgame.

People who work in government projects know how difficult it is to manage data in a public database. Since it is publicly available, they are at constant risk of attack by people who can use it to track individuals and threaten individual privacy.

Some businesses can even use this data for ulterior motives like remarketing or targeting ads. What is even worse is that the same data can be used to influence user behaviors (we all know about the Cambridge Analytica scandal, right?).

For this reason, ensuring data integrity for open data is not only a top priority but also a big challenge among all of us.

Here is how you can ensure data integrity for open data.

1. Open Data Procedures

The best way to preserve the integrity of open data is to ensure everyone takes responsibility for it. There should be strict guidelines on how to use the data available in the open database. Rules should be clearly defined, and people violating them should be held accountable.

Apart from this, people managing open data should also follow the following steps:

- Conducting purpose assessments so that everyone knows why the open data has been released
- Conducting privacy impact assessments to know the impact of releasing said data in public

All these little measures will make sure no one is able to misuse the open data, and if anyone does so, he/she can be held accountable. It will help you preserve data integrity even if the data is in an open database.

2. Using Privacy Enhancing Tools

Another way to ensure the integrity of open data is to use tools that will implement different levels of privacy into your data when you are releasing it in public. This way, you can ensure that your data is protected even before a breach and its integrity remains intact.

Here is a list of the top 10privacy-enhancing technologies that will help you safeguard your data even when it is stored in a public database.

3. Using Data Pseudonymization/Anonymization

Data Pseudonymization/Anonymization are techniques to remove/replace personally identifiable information in datasets. This way, it becomes difficult for anyone to breach someone's privacy and the data integrity remains preserved.

4. Using Data Aggregation

Data aggregation is another practice to preserve data integrity in which we group and release detailed data in the form of statistics or metadata. This way, the data in the database remains intact, and no one can get into someone's confidential details.

5. Using Access Restrictions

Using access restrictions can also be used as a method to preserve the integrity of open data. You can classify data into different levels and grant each user access according to their access level. It will give you control over how much data you want to share, even if it sits in an open database.

There are four types of access restrictions that can be implemented:

Managed: Only authorized people can access the data.

Attributed-Based: End-users must fulfill some criteria to access the data.

Public: Everyone can access the data but under some terms and conditions.

Raw: Anyone can access and use the data without restrictions.

6. Using Reuse Restrictions

You can also preserve the integrity of open data by restricting how the data is being reused. Licenses grant you the ability to control how people who can access your data can distribute it.

Doing so is helpful, especially if you are a government body and the data you are handling is too sensitive for reuse, even if you have to list it in a public database. Then, no one can misuse it and affect its integrity.

In a Nutshell

Collecting data is not a big challenge these days. Preserving data integrity is. We are creating and gathering so much data that storing and managing it has become a major issue. Chances of errors have increased multiple times, and cybercrimes have become more frequent.

Unless you have a strategy in place, preserving data integrity will be a big challenge. Hopefully, with this guide, you gained some ideas of how you can preserve data integrity and empower your decision-makers to make the right decisions.

In the end, you need to remember that healthy and integrated data can save you both a fortune and a lot of energy that you otherwise would have to spend dealing with data compromise. So, you must try your level best to preserve the integrity of data in your organization.