DATA GOVERNANCE AND PRIVACY

SkillWeed

TABLE OF CONTENTS

INTRODUCTION

Data may be defined as the information that has been translated into a form that is efficient for movement or processing for better understanding. Relative to today's computers and transmission media, data may be seen as the information converted into binary digital form. Thus, data can be defined as a representation of facts, concepts, or instructions in a dignified manner, which should be suitable for communication, interpretation, or processing by human or electronic machine. For data to be more useful and helpful, it passes through three basic stages which are; Input, Processing and Output.

INPUT

In this step, the input data is prepared in some convenient form for processing. The form will depend on the processing machine. For example, when electronic computers are used, the input data can be recorded on any one of the several types of input medium, such as magnetic disks and tapes.

PROCESSING

In this step, the input data is transformed to produce data in a more useful form. For example, pay- checks can be calculated from the time cards, or a summary of sales for the month can be calculated from the sales orders.

OUTPUT

At this stage, the result of the proceeding processing step is collected. The particular form of the output data depends on the use of the data. For example, output data may be pay-checks for employees.

In a time when organizations progressively depend on data for every aspect of their business, Data governance is a significant part of compliance. The systems will take care of the technicalities of storage, handling, and security.

But it is the responsibility of the governance organization that ensures that policies are defined, procedures are sound, technologies are appropriately managed, and data is protected. Data must be properly handled before being entered into the system, while being used, and when retrieved from the system for use or storage elsewhere.

In simple terms, Data Governance refers to setting the internal standards, and data policies that apply to how data is received, kept, processed, and disposed of. It establishes who has access to what data and what data is subject to regulation.

Compliance with external standards imposed by industry groups, government agencies, and other stakeholders is also part of Data Governance.

Thus, data governance sets the policies and procedures for establishing data accuracy. For a proper governance of data, there has to be a well organised structure:

COMPONENTS OF DATA GOVERNANCE

Data governance is the foundation of all data management programs. It is an essential sector that supports all other data management information areas like Data Warehousing, Business Analytics, Big Data, Master Data Management, and so on. Data governance has 10 major components that exist in order to meet the data management requirement. Listed as follows;

PEOPLE

This consists of the data governance professionals, data stewards and other key business and IT staffs. These are the backbone of any data governance program. They establish and develop workflows to ensure that the enterprise data governance requirements are met. Organizations need to invest heavily in training and education to ensure that this component provides the maximum amount of value to the enterprise.

DATA STRATEGY

The data governance team plays a key role in the development and implementation of the roadmap of an organization’s enterprise data strategy.

Building an enterprise data strategy is a vibrant step in the data management journey; however, many organizations fail to understand the need for a formal and documented data strategy. A data strategy is an executive document which provides the high-level, enterprise requirements for data and a strategy to ensure that those requirements will be met.

DATA PROCESS

programs need to establish key data processes for data management. Common policies include data issue tracking, data quality monitoring, data sharing, data lineage tracking, impact analysis, automated data quality testing, and many others.

DATA POLICIES

A data policy is a high-level set of one or more statements that will state expectation and expected outcomes, to influence or direct data habits at an enterprise level. Data Governance programs establish data governance policies for data management. Policies include outbound data sharing, regulatory adherence, and many others.

DATA STANDARDS AND RULES

A data standard provides a framework and all-encompassing approach for ensuring observance to a data policy. An example of a data standard could be the use of the ISO 3166 standard for the definition of the codes for the names of countries, dependent territories, special areas of geographical interest, and their principal subdivisions.

DATA SECURITY

Involves protecting digital data, such as those in a database, from damaging forces and from the unwanted actions of either from an authorized or unauthorized users, such as espionage, cyber-attack or a data breach.

COMMUNICATION

A communications plan is the key artefact of enterprise data governance communications. The plan identifies how to present data governance and stewardship challenges and successes to the various stakeholders and to the rest of the organization. The communications plan highlights the right business cases and presents their results. Data governance communications include all written, spoken, and electronic interaction with association audiences who have the obligation to know about the activities of the data governance team.

SOCIALIZATION

The data governance socialization plan is the foundation of data governance socialization. The data governance socialization plan is a plan and process by which data governance activities are incorporated into an organization's policies, internal culture, hierarchy, and processes. The plan is unique to the organization as it must be custom-made to the enterprise's culture and standards of conduct.

METRICS AND KPLS

Establishing business metrics and KPIs for monitoring and measuring the overall impact of data governance on the business program is very germane to the program's success. The metrics and KPIs must be measurable, tracked over time, and consistently measured the same way year in and year out.

TECHNOLOGY

The data governance program will need various technology to be as automated as possible.

Smaller data governance programs will typically use the technology stack which they already have within their enterprise. Larger data governance initiatives will have to purchase software that is specific to data governance relating to the functions the organization requires. Data governance software simplifies the process of scanning databases & files to capture needed metadata, management of the metadata, automating the data stewardship workflows, decision trees, social voting, collaboration, and many other data governance and stewardship functions needed.

OBJECTIVES OF DATA GOVERNANCE

Data governance was set up for a whole lot of reasons, of which some of its sole objectives are;

- Ensuring the use of standard, repeatable processes for data entry and reporting.
- To support a culture of informed decision making based on clean, consistent and understandable data.
- The fostering of an organized system to manage data effectively and ensure clean, consistent data.
- One of the main objectives of data governance is to create and maintain communication network for changes to Excellent, data entry criterions or business processes.
- To create and maintain business development documentation.
- Another aim and objective of data governance is to protect the privacy, integrity and security of all data through an intended and guided process.

Another data governance goal is to ensure that data is used properly, both to avoid hosting data errors into systems and to come against potential misuse of personal data about customers/clients and other sensitive and delicate information. This can be accomplished by creating unchanging policies on the use of data, along with the procedures to monitor the usage of data and implement the policies on a constant basis. In addition, data governance can help to strike a balance between data collection practices and privacy mandates.

IMPORTANCE OF DATA GOVERNANCE

Without effective data governance, data discrepancies in different systems across an organization might remain a huge problem. This could bring about complications in data integration efforts and create data reliability issues that affect the precision of business intelligence.

Other than a more accurate analysis and stronger regulatory compliance, another benefit/importance of data governance is that data governance provides an upgraded data quality, with a lower data management costs, it also provides an increased access to needed data for data researchers, other analysts and business users. Ultimately, data governance can help improve business decision-making by giving executives better information. Ideally, that will lead to competitive advantages and increased revenue and profits in the business organization.

DECISION MAKING

One of the crucial importance of data governance is that it helps in making a better decision.

This applies to both the decision-making process, as well as the decisions themselves. A well-governed data is more helpful, making it easier for necessary users to operate professionally. It also means decisions will be based on the right data, ensuring greater accuracy and trust.

EFFICIENCY

Good data is highly valuable in the era of data-driven business. Therefore, it should be treated as the asset it is. Consider a manufacturing business' physical assets, for example. Well-run manufacturing businesses ensure their production-line machinery undergoes regular inspections, maintenance and upgrades so the line operates smoothly with limited down-time. The same approach should apply to data.

IMPROVES BETTER UNDERSTANDING

Data governance helps improves a better understanding on what each data is all about and where it is stored. When implemented well, governance provides a comprehensive view of all data assets. It also brings about proper accountability. By assigning permissions, it is way easier to determine who is responsible for specific data.

HIGH DATA QUALITY

As data governance aids in discoverability, businesses with an effective data governance programs also profit from an improved data quality. Although technically two separate initiatives, some of their goals overlap. These include, but are not limited to, the standardization of data and its consistency. One way to clearly differentiate the two programs is to consider the questions posed by each field. Data quality wants to know how useful and complete data is, whereas data governance wants to know where the data is and who is responsible for it. Governance improves data quality, because answering the latter makes it easier to tackle the former.

HIGH SECURITY

One of the main goals of any data governance plan is security. This includes outlining and verifying the requirements for data distribution policies, but also maintaining observance against any external cyber-attack, internal equipment failure or system crash that could compromise sensitive data. This should all go into a business continuousness plan, which no company should be without.

DATA GOVERNANCE PRINCIPLES

These data governance principles apply to any data governance strategy and they act as rules for effective data management.

TRANSPARENCY

Transparency consists of building a framework in which information flows clearly to the team members so that they're aware of any changes at all. In data governance, there should also be a reference line to measure against. Without a baseline, any reference is not affixed to context and therefore cannot be of any use in terms of transparency or coming up with improvements. A good way to think of data governance is to liken it to quality control, both of which can play an important role in total quality management.

DATA STEWARDSHIP

The data steward is mainly responsible for making sure that the quality of data remains high, which makes it accurate, accessible, consistent, complete and updated. The data steward does not necessarily have to be an individual, but a team allocated with the task of maintaining the data governance. The team consists of database administrators, business analysts and other personnel who understand the context of data within the organization. The data steward works with people who manage the overall data life cycle to make sure the data toe the line of the organization's data governance policies.

ACCOUNTABILTY

To guarantee that the data is being governed to support business goals there must be accountability. Data governance systems don't just turn on and work without oversight. There must be data owners and data stewards designated to manage, monitor and report on the quality of information. Each sectors must take responsibility and make an account on all of their operations.

DATA QUALITY MAINTANACE

Maintaining the data quality of any data governance plan requires data editors, data mining tools, data differencing utilities, data linking tools, workflow as well as project management tools. Data quality is the pioneer for most data governance tasks.

Quality in this sense means the accuracy, completeness and consistency over the whole data structure of the organization. Part of data quality is data scrubbing or data cleansing, which identifies, relates and removes any form of duplicated data.

DATA GOVERNANCE STRUCTURE

The structure should include the following;

DATA SCOPE

This consists of the Data analyst and the operator, it is an electronic display that shows the content of the information being transmitted over a communications channel.

ORGANIZATIONAL CONDUCT

For a proper governance of data, there must be an exceptional diligence and consistency in the roles and responsibilities between accountable owner, head of data, IT, business team, and executive sponsor.

DATA GOVERNANCE POLICY

This is a documented set of guidelines for ensuring that an organization's data and information assets are managed consistently and used properly. Such guidelines typically include individual policies for data quality, access, security, privacy and usage, and they specify different roles and responsibilities for implementing those policies and monitoring compliance with them.

OVERSIGHT

This is really important and necessary, in order to provide a framework that ensures data are captured accurately and consistently from a variety of sources, and to create data policies that maximize access to this data to all appropriate personnel.

The data governance operates under some certain guidelines and procedures which are:

PROCEDURES AND DOCUMENTATION

More than just a requirement to keep auditors satisfied, documentation must clearly outline all processes. And procedures should also be reinforced through training and with motivational incentives.

DATA INTEGRITY

Considerations for data integrity must be built into procedures according to the data governance model and framework.

AUDITS AND QUALITY CONTROL

Build periodic checks of data validity into all procedures to verify processes and procedural compliance. A regular schedule of checks by a quality team works best in this process.

DATA GOVERNANCE CHALLENGES

One of the biggest challenges of data governance can be organizational and personnel issues. Every business transformation requires accountable roles and responsibilities with a champion to lead the change. It also requires a culture shift from viewing data management as a boring, low-level job to one of extreme importance. If employees touch data (especially critical data) and if they create it, change it, use it, or move it around in some way, they need to understand the role they play in properly maintaining that data and take accountability.

Another big challenge is the rapid proliferation of data that is only becoming more rampant over time. Much of this new data is either unstructured or different from what we have seen or worked with in the past. This not only taxes existing systems and databases but brings the need for new procedures and additional requirements for governance.

DATA GOVERNANCE TOOLS AND TECHNOLOGY

Information steward applications (ISA): This assist in data profiling and monitoring the performance of the enterprise's data governance policy. It facilitates executing information governance initiatives across the business units, enforcing quality standards with data validation, and measuring the improvement of data quality processes.

Metadata management solutions, often referred to as EMM (enterprise metadata management), categorize and consistently organize an enterprise's information assets and has become increasing important in the time of Big Data. Information of the data asset that is maintained include type, tags, source, and dates.

Augmented data management, or augmented data integration, enhances existing enterprise data with information attained using new technologies such as

AI (artificial intelligence) and machine learning. The goal is to improve decision making and help some applications in becoming more self-tuned.

Information lifecycle and content management technologies control data volumes and manage risk with automated information archive, retention, and destruction policies. Content management-specific capabilities can also streamline business processes by digitizing documents and integrating relevant content with transactions and workflows.

Some of the most used data governance tools includes the following;

ALATION DATA CATALOG

The Alation Data Catalog offers users a single point of reference for numerous data sources. Thus; making it easier to identify and locate the information they want. It is one of those tools which helps in the mechanisation of Governance operations such as updating data dictionaries and training users about good Governance principles, as well as offering collaborative options for sharing information.

ATACCAMA

This is a Platform that provides services for a self-driven Data Management and Governance. It makes data administration modest with AI-driven automated abilities. With a "self-driving" strategy aimed to mechanise, increase efficiency, and ease of use, Ataccama automatically analyses data quality and classifies data in order to help firms prioritize and concentrate. All essential data assets can be placed under security and privacy regulations thereby making data available only to those who are in need of it.

COLLIBRA

Collibra tool is a Data governance tools for enterprises. It automates data operations and keeps cross-functional teams on the same page. It also provides a Natural Language Search, Data Governance Automation, and Data Stewardship for the benefits of the users. With the help of this too, users can also have an interactive data lineage diagrams to graphically understand data elements like regulations, problems, relationships, and flows.

OVALEDGE

This tool is a data catalogue tool and a cost-effective Data Governance solution. Companies can adapt the tool to meet their business demands due to its agile architecture. It also makes of algorithms and user inputs to illustrate relationships between your data, giving you a 360 degree of your data.

ATLAN

Atlan's data workspace platform supports Data cataloguing, Data Quality, Data Lineage, and Governance. For developing a shared understanding of data, the software includes a Google-like Search interface, automated Data Profiling, and a searchable Business Wordlist. No matter where your data flows, users can manage bandwidth utilization and adoption throughout an ecosystem using granular Governance and access restrictions with the help of this tool.

INFORMATICA

This data governance tool creates a data catalogue by scanning across many Cloud platforms automatically. It allows the tracking of data lineage to be done automatically. This tool also keeps track of data migration, from high-level system views to granular column-level lineage, and analyses the impact. Informatica disassembles silos and brings together IT, security, and business teams to guarantee that data is compliant with regulations such as **GDPR** (General Data Protection Legislation).

SAP MASTER DATA GOVERNANCE

SAP Master Data Governance software (SAP MDG), is accessible on- premises or in the Cloud. This particular tool is developed to perform effective enterprise data management, minimize risk, improve compliance, and also to lower the total cost of ownership. SAP MDG consolidates master data and automates replication and syndication of master data throughout the system landscape. It also measures, monitors, and improves master data quality and workflows in the operations.

ERWIN

Erwin combines Data Governance, Enterprise Architecture, Business Process, and Data modeling and narrow it into one software platform. By linking physical metadata to particular business terminology, it allows customers to locate and harvest data, as well as arrange and deploy data sources. Erwin can assess complicated lineages spanning systems and use cases by importing information from data integration tools and cloud-based platforms.

IBM DATA GOVERNANCE

IBM Data Governance tool offers features such as a non-rigid Data Governance plan, Data Cataloging, and acquiring valuable data for Big Data creativities. It also enables privacy and security features, such as the ability to secure personally distinguishable information, predictive consumer intelligence, and personal health data.

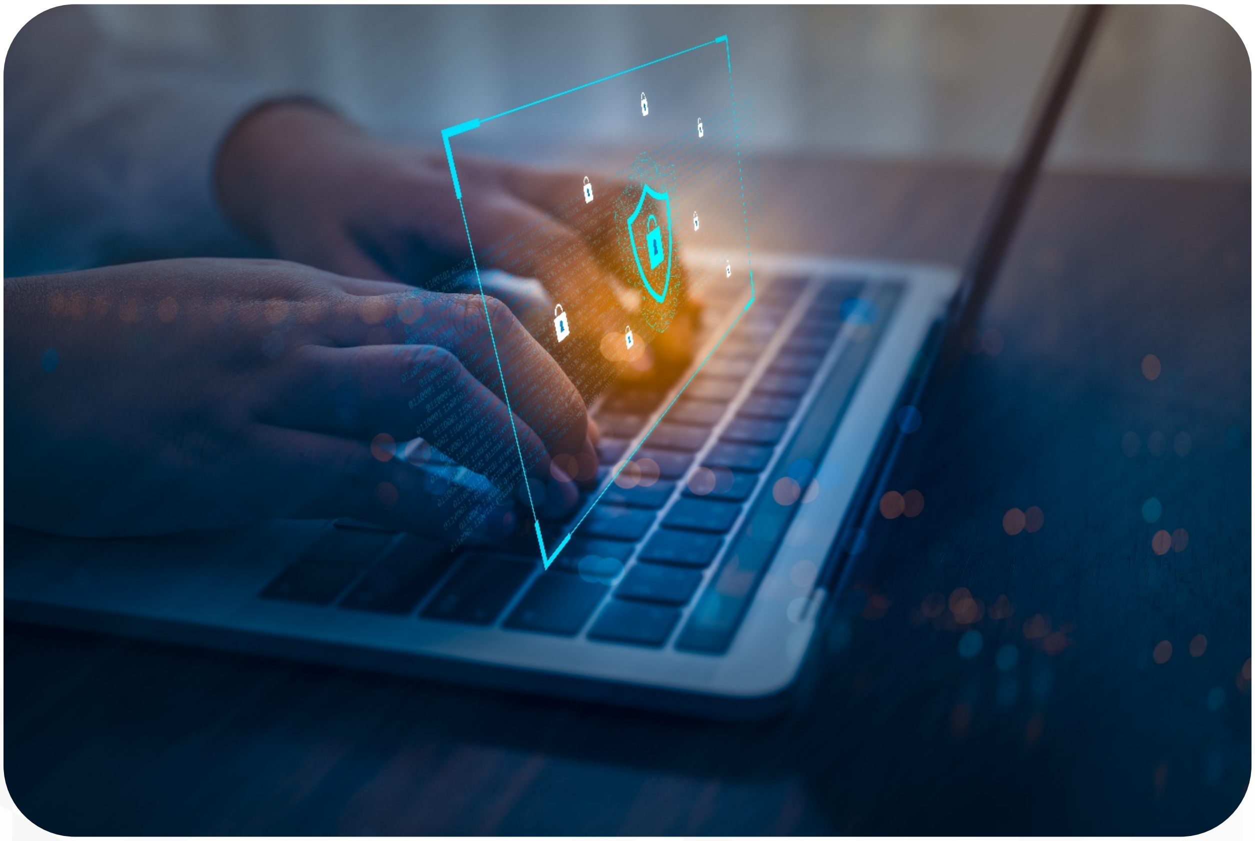
APACHE ATLAS

Apache Atlas is amongst the best data governance tools. It is an open-source tool that is built upon a foundational set of metadata control and information governance talents for businesses with information-extensive platforms.

ALEX SOLUTIONS

Alex’s technology-agnostic data catalogue is available as a full Cloud or hybrid solution. It can harvest and ingest metadata from a variety of platforms, both on-premises and in the cloud, and supports connectors for all major cloud providers. Alex’s automatic data lineage competency allows users to quickly alter the context between business, user, application, and technology lineages easily.

BENEFITS OF DATA GOVERNANCE TOOLS



As data and apps have grown more important to businesses, Data Governance Tools and Technologies have become more important to ensure the reliability of data assets. Below are some few key points of the relevance of data governance tools;

EFFECTIVELY COLLABORATE BUSINESS INFORMATION

Data Governance Tools organize all of a company's data into an easy catalogue to navigate, allowing people with the appropriate rights to view and collaborate on data in minutes. This is accomplished through self-service, thereby reducing the workload on data teams.

COMPLIANCE AND REGULATION

Every sector has its own set of rules and regulations to follow, and an act of disobedience to these regulations may result in weighty penalties. Hence, Data Governance Tools ensure compliance with regulations while also safeguarding security and privacy.

CREATES A STANDARDIZED VIEW

In business terminology and concepts, inconsistencies are common as various users employ different phrases to describe the same concept.

Hence, a business dictionary is included in Data Governance Tools to guarantee that all terminology is standardized and that everyone is on the same page without misunderstanding.

HIGH DATA LITERACY AND UNDERSTANDING

Everyone in your organization must be data literate to leverage data assets to make better business choices if you want to succeed as a data-driven firm. A Data Governance Tool is a simple platform that allows corporate users to locate and collaborate on data assets. As practically every operation can be completed through self-service, this allows users to become more literate while also relieving the burden on data teams.

GENERAL DATA PROTECTION REGULATION (GDPR)

GDPR can be considered as the world's strongest set of data protection rules, which enhance how people can access information about them and places limits on what organisations can do with personal data. GDPR came into force on May 25, 2018. Countries within Europe were given the ability to make their own small changes to suit their own needs. Within the UK, this flexibility led to the creation of the Data Protection Act (2018), which superseded the previous 1998 Data Protection Act.

WHAT DOES GDPR OPERATES ON?

GDPR majorly operates on giving protection to personal data. Under this, there's also a few special classifications of delicate personal data that are given greater protections. This personal data includes information about racial or ethnic origin, political opinions, and religious beliefs, membership of trade unions, genetic and biometric data, protection rightson and data around a person's sex life or orientation.

At the heart of GDPR is personal data. Broadly, this is information that allows a living person to be directly, or indirectly, identified from data that's available. This can be something obvious, such as a person's name, location data, or a clear online username, or it can be something that may be less instantly obvious: IP addresses and cookie identifiers can be considered as personal data. The crucial thing about what constitutes personal data is that it allows a person to be identified and recognised. Pseudonymised data can also be classified under the definition of personal data. Personal data is so important under GDPR because individuals, organisations, and companies that are either 'controllers' or 'processors' of it are all regulated by the law.

Although coming from the EU, GDPR can also apply to businesses that are based outside the region. If a business in the US, for instance, does business in the EU then GDPR can apply and also if it is a controller of EU citizens.

THE PURPOSE AND SCOPE OF GDPR

One of the sole purposes of the General Data Protection Regulation (GDPR) is to protect individuals' fundamental rights and freedoms, particularly their right to protection of their personal data and privacy. A further purpose of the General Data Protection Regulation is to create a constant and synchronized level for the protection of personal data within the EU so that the free movement of personal data within the Union is not held up. This is achieved through the regulation being directly applicable in the various member states and through the same rules applying throughout the Union.

The General Data Protection Regulation applies to personal data processing linked to the EU, either when the entity processing the personal data is established within the EU or when an entity outside the EU offers goods and services to people within the Union or monitors their behaviour here.

GDPR FINES AND PENALTIES OF VIOLATION OF POLICY

National authorities can or must assess fines for specific data protection violations in accordance with the General Data Protection Regulation. The fines are applied in addition to or instead of further remedies or corrective powers, such as the order to end a violation, an instruction to adjust the data processing to comply with the GDPR, as well as the power to impose a temporary or definitive limitation including a ban on data processing. For the provisions which relate to processors, he may be subject to sanctions directly and/or in conjunction with the controller.

The fines and punishment for each offence must be effective, proportional and dissuasive for each individual case. For the decision of whether and what level of penalty can be assessed, the authorities have a legal catalogue of criteria which it must consider for their decision. Among other things, intentional infringement, a failure to take measures to mitigate the damage which occurred, or lack of collaboration with authorities can increase the penalties. A punishable situation in a company can be revealed through practical inspection activities conducted by the data protection authorities, by an unsatisfied employee or by customers or potential customers who complain to the authorities by the press in general, especially through investigative journalism. Thus, any form of irregularities and violation of the general policies or rules and regulation of the GDPR will be severely punished.

THE SEVEN PRINCIPLES OF GDPR

The seven principles of GDPR gives a brief guide on managing data privacy and constitutes some of the practices organizations need to implement to minimize risks pertaining to data protection and maintaining compliance with the rights of the individual. These principles include:

LAWFULNESS AND TRANSPARENCY

Organizations should totally transparent on their resolution behind collecting the data and the reasons for which it is being used for. This shows that all the data processing activities are done in a legitimate manner and gives data subjects the assurance of their personal data being collected lawfully and on a pure basis.

FAIRNESS

This means the data should be handled in a way that is familiar to the individual. Any information that will be processed by the organization must be only what the individual gives a go ahead on.

PURPOSE RESTRICTION

The main aim for this principle is to limit the amount of personal data being processed. The data should be collected and managed only along the lines that is rational and justifiable to the data subject in concern.

DATA MINIMISATION, ACCURACY AND STORAGE

LIMITATION

These three principles, which is collectively known as ‘data principles’, defines the standards in which data should be handled. These principles can be found within Article 5(1) (c), (d) and (e) of GDPR respectively.

DATAMINIMISATION

This principle insists on the fact that data collection should be strictly restricted to the minimum amount of data needed to carry out a specific purpose.

ACCURACY

This enables individuals the privilege to correct any inaccuracies in their data. Organizations should have some sort of system in place to ensure constant accuracy by having periodic updates or rectifying any measures that brings about any form of inaccuracy in the operation.

STORAGELIMITATION

Organisations should have a well written policies that properly defines data disposal and preservation periods.

CONFIDENTIALITY AND INTEGRITY

A highly adequate amount of security measures should be in place to ensure protection of data against any unlawful disclosure, alteration or removal. Both technical and operational measures should be well-thought-out while guarding information within the systems and networks.

ACCOUNTABILITY

This principle essentially breaks down into these two factors; that the organization should be held accountable to complying with all its pertinent requirements of GDPR and should be able to demonstrate its compliance in a clear and concise manner. This principle helps build the trust of the data subjects and improve your reputation as an organization that takes data privacy seriously. There should be documents in place stating how personal data will be processed, its purpose and time period in which the data will be processed for.

DATA AUTONOMY

Delicate data should be protected in accordance with the laws and regulations guiding the data of the geographical location of the data subject. This requirement came into prominence and lime light with the rise of cloud-based applications and the increased amount of data transfer across international borders.

RIGHTS OF DATA SUBJECTS

Individuals, likewise, have the right to some certain things about their data before the organization can make any move pertaining to such data. These rights include;

ACCESS AND MODIFICATION

One of the rights of individuals is that they have the right to submit subject access requests and be duly informed by the organization whether or not their information is being processed or not. When getting the request, the organization must send a copy of the personal data they collect about the individual. The right to modification gives individuals the freedom and civil liberties to reach out to organizations to correct any inaccuracies and errors in the information they have about them.

RIGHT TO BE INFORMED

This gives individuals the right to be informed about the information being collected about them, its purpose, the processes and systems used in which the data is processed and the retaining period which states how long the data will be kept for. This also provides them with the right to file a complaint if necessary.

DATA PORTABILITY

This gives individuals with the right to obtain their own copy of personal data previously collected by the organization. The data subject can also send in a request to have the information transferred to another organization. This can only be applicable if the individual has come to an arrangement with the organization by the means of a contract or consent and the processing activities are carried out through programmed means.

RIGHT TO ERASE

Individuals have the right to request the deletion of their personal data for some unsatisfactory reasons such as; if the processing is done unlawfully, if the data is no longer required or if the individual withdraws the consent. An also, a clearing of data can be done if the erasure is required in order to be acquiescent with a legal obligation.

RIGHT TO RESTRICT PROCESSING

This gives individuals the right to refrain organizations from processing their data. This right can only be exercised on some certain grounds such as; if the data is processed unlawfully, if the data is no longer needed by the controller, if the data is incorrect and also if the processing activities have been put on hold to verify a data erasure request. On these grounds, an individual have the right to restrict any processing of their data by the organisation.

RIGHT TO AMENDMENT

Individuals also have the right to ask organizations to correct any inaccurate information organizations may have about them. Once the request is sent, organizations are given a month's time to respond to it and make an amendment on the situation.

RIGHT TO OBJECT

This gives individuals the right to object the processing of personal data for marketing or other legal purposes.

DATA PRIVACY

Data Privacy describes the practices which ensure the data shared by customers is only used for its intended purpose. In a world with ever-growing amounts of data, privacy is a crucial topic to analyse. Data privacy, which is sometimes also referred to as information privacy, is an area of data protection which is concerned with the proper handling of sensitive data including

notable personal data but also other confidential data, such as certain financial data and intellectual property data, to meet regulatory requirements as well as protecting the confidentiality and immutability of the data. Privacy varies from one geographical location to another in meaning, In the European Union, privacy is recognised as an absolute fundamental right and in some parts of the world privacy has often been regarded as an element of liberty, the right to be free from intrusions by the state. In most geographies, privacy is a legal concept and not a technology, and so it is the term data protection that deals with the technical framework of keeping the data secure and available.

In other words, Data privacy is the branch of data management that deals with handling personal data in compliance with data protection laws, regulations, and general privacy best practices. Data privacy needs to be a top priority for businesses. Failure to comply with data privacy regulations can lead to big losses. Think legal action, steep financial penalties, and brand damage.

IMPORTANCE OF DATA PRIVACY

Keeping private data and sensitive information safe is paramount. If items like financial data, healthcare information, and other personal consumer or user data get into the wrong hands, it can create a dangerous situation. The lack of access control regarding personal information may expose an individual to fraud and identity theft. In a business atmosphere, the securement of data is a very important act. Few importance are as follows:

REGULATORY COMPLIANCE

Managing data to ensure regulatory compliance is arguably even more of an importance. A business may have to meet legal responsibilities about how they collect, store, and process personal data, and non-compliance could lead to a huge fine. If the business becomes the victim to a hack or ransom ware, the consequences in terms of lost revenue and lost customer trust could be even worse.

BUSINESS ASSET MANAGEMENT

Data is perhaps the most important asset a business owns. We live in a data economy where companies find enormous value in collecting, sharing and using data about customers or users, especially from social media. Transparency in how businesses request consent to keep personal data, abide by their privacy policies, and manage the data that they've collected, is vital to building trust with customers who naturally expect privacy as a human right.

Additionally, a data breach at the government level may risk the security of entire countries. And if one occurs within your company, it could make your proprietary data accessible to a competitor. A data breach is an intentional or unintentional release of confidential data that exposes it to an untrusted environment. Other common terms for this include "unintentional information disclosure," "information leakage," "data leak" and "data spill." Data breaches can occur in a variety of ways and contexts, from malicious attacks by criminal hackers, political activists, or foreign governments, to careless processing when disposing of computer equipment or other data storage media.

STEPS TO EFFECTIVE DATA PRIVACY

DEFINE SENSITIVE DATA

Sensitive data is any data that, if lost, stolen, or exposed, could financially hurt your organization, cause reputational damage, or harm the data owner. The first step in creating a data protection program is to determine which information your organization collects meets the definition of sensitive. This will clarify exactly which data needs to be protected and the legal regulations that cover it.

FULLY UNDERSTAND THE DATA LIFECYCLE

To protect your sensitive data most effectively, you need to understand its lifecycle. The data lifecycle stages include create, store, use, share, archive, and destroy. Knowing the stage of each piece of sensitive data determines in large part which policies and tools you should implement to best protect it at each point of its lifecycle.

BE AWARE OF EACH SENSITIVE DATA REGULATION YOU'RE SUBJECTED TO

Compliance is the other major factor influencing the policies and tools you implement to protect your organization's data. For example, storage practices must include encryption and firewalls to comply with data privacy regulations. They also call for access controls and audit logs to trace data use and sharing back to an individual. Lastly, regulations often require data to be disposed of in a timely and secure manner, so policies need to be implemented to ensure compliance.

DECIDE WHO CAN HAVE ACCESS TO YOUR INFORMATION

Authentication methods can include passwords, PINs, access cards, or biometrics, such as fingerprints or facial recognition. Having authentication in place to access certain data will help IT departments keep track of any changes made to it and trace those changes back to a specific person. Access to sensitive data should only be given to employees needing it to fulfil their job responsibilities. To ensure this, require authentication and authorization permissions to access certain data. All authenticated individuals should have permission roles assigned to them. Not everyone needs modification abilities, and only those requiring this access should be allowed. Assigning roles such as viewer, editor, and administrator can help limit opportunities for sensitive data misuse.

INVOLVE ALL EMPLOYEES IN SECURITY ALERTNESS

It's essential that your organization educates all individuals, even those who don't touch any sensitive data, about the data security responsibilities attached to certain roles. Everyone should understand that their actions regarding sensitive data can directly affect the organization's success and reputation, as this will help employees recognize and call out improper handling of sensitive data, as well as prevent any inadvertent sharing of it.

CONDUCT A FREQUENT CHECK-UP

In addition to fortifying your data's storage locations, be prepared to back up that data as often as needed and have different, yet just as secure, places available to store it. For example, if your primary storage is cloud-based, consider backing up to a physical location. In the case of a breach, you can use these backups to restore lost or corrupted data, which can ultimately lessen the financial blow to your organization.

DOCUMENT ANY PROCESSES USING SENSITIVE DATA

Many data privacy regulations require you to be able to share with consumers how their sensitive data is being used in your organization's business processes. By documenting the types of data collected, contexts of use, and collection, storage, and sharing methods, you uphold compliance while also gaining a clearer picture of the data you possess and how it's handled. In the unfortunate case of a compromise, you can audit this documentation to identify where in your organization's process or infrastructure or with whom a vulnerability resides.

TAKE ACCOUNT OF YOUR DATA

Everything, from security to compliance, begins with locating your sensitive data. To find it, look at cloud repositories, physical file servers, computer hard drives, HR databases, **CMDB** or **eGRC** platform, and any other system of record. Once you identify sensitive data, you know exactly what to protect to uphold compliance and reduce the risk of data breaches.

You're able to apply increased security measures for all existing data at the various stages of its lifecycle and will be better prepared to handle the creation of new data moving forward.

ORGANIZE THE DATA YOU WISH TO PROTECT

To protect data and meet compliance requirements, you must classify data according to its level of sensitivity. Classification systems help you set those use and modification access controls we mentioned earlier, acting as a natural next step to protect data once discovery is complete. Classification schemes you can use include role-based, data-oriented, access- or location-based, and hybrid. Most organizations categorize or bucket data as variations of a four-level data classification schema; public, private, confidential, and restricted.

AUTOMATE PROCESSES FOR STRONG ONGOING PROTECTION

If all of this sounds difficult to do manually, that's because it is. Human error is inevitable, which can cause an oversight during manual discovery and leave data unprotected.

Manual classification can lead to inconsistent labelling or overlooking the critical context of a piece of data, which eventually causes it to be misclassified and left vulnerable.

Once your program is in motion, automation tools are what will help it run efficiently and accurately. They enable enterprise organizations to gain critical visibility of their must-protect assets across clouds, networks, devices, and endpoints. You can't have an effective data protection program without automation software on your side.

DATA PRIVACY VS DATA SECURITY

Data privacy is related to, but not the same as, data security. They do have some overlapping obligations:

- **Access Control**: Preventing unauthorized access to and use of data is the cornerstone of privacy, and possible only through security.
- **Integrity of Data**: Making sure that data is accurate and not altered is both a privacy and a security concern.
- **Accountability**: Company policies relating to data should document both privacy and security.

But privacy and security have different emphases. Data security is concerned with ensuring the confidentiality, integrity, and availability of all data. Security professionals implement cyber security measures such as authorization and data encryption.

They prevent data breaches and defend against malicious attacks.

On the other hand, Data privacy, by contrast, focuses on information about individuals. Privacy rules determine what types of PII may be collected, about whom, to what extent, and what can be done with it. Businesses must ensure that only the appropriate access rights are granted to people in the organization, to partners with which they share data, and to the general public.

Data privacy focuses on how personal data is collected, used, and shared, in other words, its governance. Regulations and laws addressing data privacy can vary by state and country in terms of how stringent they are and how they are enforced. Worldwide, countries are coming to the realization that the strict guidelines designed to protect personal data privacy are in the best interest of both an organization and individuals. The European Union's General Data Protection Regulation (GDPR) is the strictest regulation to date, with other countries modeling regulations after the privacy mandates of the GDPR. Some of note: the California Consumer Privacy Act (CCPA), Brazil's Lei Geral de Protecao de Dados (LGPD) and Canada's proposed Digital Charter Implementation Act, to name a few. While these enacted and proposed regulations are a huge step in ensuring data privacy, without a solid data security foundation and technological solutions in place, data privacy simply cannot happen.

Data security, as opposed to data privacy, focuses on how data is protected from the many external and internal threats that exist. Data security policies and procedures can mitigate cyber threats and inadvertent misuse; however, just putting these measures in place does not typically fully address data privacy concerns and regulations. Data security encompasses the actual solutions an organization puts in place to protect digital data at all points. From endpoints to networks to the perimeter. A comprehensive data security policy should form the blueprint for your data security measures and cover three key areas: people, processes, and technological solutions to help enforce any policies set to surround and protect sensitive and private data.

DIFFERENCE BETWEEN DATA PRIVACY AND DATA SECURITY

First, data privacy is NOT the same as data security. Data security is all the measures, policies, and technologies taken to protect data from external and internal threats. However, applying data security measures alone does not necessarily satisfy data privacy requirements. Data privacy still requires adherence to regulations surrounding how the data organizations secure is collected, shared, and used. Data security protects data from malicious threats; data privacy addresses responsible governance or use of that data.

When developing data security policies, the focus of protection measures is on preventing unauthorized access to data. Tools such as encryption, user authentication, and tokenization can all amp up an organization's security stance. When tackling data privacy concerns, the focus is on data being procured, processed, stored and sent in compliance and with consent of the data subject. If an organization is gathering data, individuals need to know what type of data will be collected, why it is needed and who will share this data for transparency.

In addition, the data subject needs to agree to these terms.

Using data with respect to an individual's privacy is the key to data privacy. Data security measures can help ensure that personal identification in collected data is protected.

Data privacy or Information privacy is concerned with proper handling, processing, storage and usage of personal information. It is all about the rights of individuals with respect to their personal information. While Security applies to all types of information, whether it's PII or not. The question of whether information is personally identifying simply determines the level of security necessary. PII requires the highest security standard. However, privacy encompasses a wider set of obligations than security, including:

- **Data lifecycle**: The data lifecycle for PII must begin with a clear purpose for collecting user data. It also maps how PII is managed, from collection to deletion.
- **Data ethics**: Ethics extend beyond lawfulness and compliance with data privacy regulations. Ethical behaviour towards personal data includes transparency, openness, and fairness regarding how that data is handled.

- **Data quality**: While ensuring the accuracy of user data isn't solely the responsibility of data privacy professionals, it is vital to maintaining data privacy. For example, if patient records aren't up to date, test results could go to the wrong person.

Data security is focused on protecting personal data from any unapproved third-party access or malicious attacks and exploitation of data. It is set up to protect personal data using different methods and techniques to ensure data privacy. Data security ensures the integrity of the data, meaning data is accurate, reliable, and available to authorized parties. Data security methods includes;

- Activity monitoring
- Network security
- Access control
- Breach response
- Encryption
- Multi-factor authentication

In short, data privacy and data security are, by no means, the same terms. Data privacy is about proper usage, collection, retention, deletion, and storage of data. Data security is policies, methods, and means to secure personal data.

THREE PILLARS OF INFORMATION SECURITY

Confidentiality: This prevents sensitive information from reaching wrong people, while making sure that the right people can use it.

Integrity: maintains the consistency, accuracy, and trustworthiness of information over its lifecycle.

Availability: ensures that the information is available when it is needed.

These are very often referred to as the C-I-A triad, and they all have to be addressed in order to achieve a satisfactory level of information security.

The risk assessment will then be cross-referenced with the organization’s risk acceptance criteria (these are developed in line with the organization’s risk appetite, i.e. their willingness to accept a predefined level of risk) and consequently, a risk treatment plan can be developed. Like many things in life where nothing is perfect, the same goes for security; there is no such thing as a 100% secure system. There are only acceptable levels of risk.

DATA PROTECTION AND RIGHT TO PRIVACY

Data protection is a legal mechanism that ensures privacy. Privacy is usually defined as the right of any citizen to control their own personal information and to decide about it (to disclose information or not). Privacy is a fundamental human right. The July 2015 appointment of the first UN Special Rapporteur on the Right to Privacy in the Digital Age reflects the rising importance of privacy in global digital policy, and the recognition of the need to address privacy rights issues the global, as well as national levels.

FRAMEWORKS FOR SAFEGUARDING THE RIGHT TO PRIVACY AND DATA PROTECTION:

The International Covenant on Civil and Political Rights (ICCPR) is the main global legal instrument for the protection of privacy. At a regional level, the main instruments on privacy and data protection in Europe is the Council of Europe (CoE) Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data of 1981. Although it was adopted by a regional organisation (CoE), it is open for accession by non-European states. Since the Convention is technology neutral, it has withstood the test of time. The EU Data Protection Directive (Directive 95/46/EC) has also formed an important legislative framework for the processing of personal data in the EU and has had a vast impact on the development of national legislation not only in Europe but also globally. This regulation has also entered a reform process in order to cope with the new developments and to ensure an effective privacy protection in the current technological environment.

The collection of data and information is being regulated in a way that protects the rights of individuals. Such that;

1. The collection and use of personal data should be limited to purposes: (1) which are stated in law and thus can be known (at least in theory) to the individual at the time of the data collection; or (2) for which the individual is aware of.
2. The data collected must be proportionate to the purpose of the ID system in order to avoid unnecessary data collection and “function creep,” both of which can create privacy risks. This is often articulated as requiring that only the “minimum necessary” data—including transaction metadata—should be collected to fulfil the intended purpose.
3. The collection and use of personal data should be done on a lawful basis, e.g., involving consent, contractual necessity, compliance with legal obligation, protection of vital interests, public interest and/or legitimate interest.
4. The collection and use of personal data should be done fairly and transparently.
5. Personal data should be accurate and up-to-date, and inaccuracies should be expediently corrected.

6. Requirements to use technologies that protect privacy (e.g., the tokenization of unique identity numbers) by eliminating or reducing the collection of personal data, preventing unnecessary or undesired processing of personal data, and facilitating compliance with data protection rules.
7. Personal data, including transaction metadata, should not be kept longer than is necessary for the purposes for which it is collected and processed. With respect to transaction metadata, people can be given an option for how long such data are retained.
8. The processing of personal data in accordance with the above principles should be monitored by an appropriate, independent oversight authority, and by data subjects themselves.

In general, personal information should be lawfully obtained (usually through freely given consent) for a specific purpose, and not be used for unauthorized surveillance or profiling by governments or third parties or used for unconnected purposes without consent (unless otherwise required under the law).

Finally, users should have certain rights over data about them, including the ability to obtain and correct erroneous data about them, and to have mechanisms to seek redress to secure these rights, all of these requirements gives boost in the support of individual rights.

IN CONCLUSION

Data governance is a data management concept. It is a measure of the control an organization has over its data. This control can be achieved through high-quality data, visibility on data pipelines, actionable rights management, and clear accountability. Data governance encompasses the people, processes, and tools required to create consistent and proper handling of a company's data. By consistent and proper handling of data, I mean ensure availability, usability, and consistency, understand ability, data integrity, and data security.

The most comprehensive governance model, say, for a global bank, will have a robust data-governance council (often with C-suite leaders involved) to drive it; a high degree of automation with metadata recorded in an enterprise dictionary or data catalogue; data lineage traced back to the source for many data elements; and a broader domain scope with ongoing prioritization as enterprise needs shift.

DATA GOVERNANCE HAS A DIRECT BUSINESS IMPACT

Data governance isn't just that rusty process that companies have to deploy in order to comply to regulation. Of course, part of it is a legal obligation, but clean governance strategy can have key business outcomes.

MAIN GOALS AND OBJECTIVE OF DATA GOVERNANCE PROGRAMME BUSINESS IMPACT:

- Maximizing the income generation potential of data
- Optimize staff effectiveness
- Enable better planning by supervisory staff **LONG TERM HEALTH:**
- Increasing consistency and confidence in decision making
- Designating accountability for information quality
- Minimizing or eliminating re-work **CONTAINS RISKS**:
- Decreasing the risks of regulatory fines
- Improving data security, also defining and verifying the requirement for data distribution policies.

The head of Data, CDO, and Data stewards are in charge of data governance;

In most organizations, data stewards are in charge to implement a framework to ensure key governance standards are met. This framework supports a set of rules and responsibilities such as assigning owners to data assets, enforcing the security of the analytics systems, adding access rights and security roles to data analysts and engineers. The framework and policies can change from one company to another. Heads of data or CDOs, that manage data analytics teams, oversee the efforts of the data stewards. They set up a clear program or strategy to prioritize the work, set standards, and define clear roles and responsibilities during a monthly or yearly committee. Data stewards support the strategy and implement the processes established by the head of data or Chief Data Officer. Good practices often focus on using a specialized governance tool, such as Castor.

The benefits of the work of the data stewards impact data governance and the efficiency of the analytics teams. It helps improve the quality of the decision making and brings visibility. They support analytics teams by maintaining high data quality standards, owners, roles to ensure smooth decisions and increasing security.

HOW TO SET UP A GOOD DATA GOVERNANCE AND PRIVACY STRATEGY

Data Architecture: Before even talking about data governance framework, a company needs the basis: a good infrastructure to begin with. Based on business needs and the company's data maturity, the nature of the data architecture framework can change a lot. Regarding storage, do you go for: on-premise or cloud? Data warehouse or Data Lake? Regarding modeling: Spark or DBT? In data warehouse or in BI tool? Real-time or batch? Regarding visualization: do you allow anyone to build dashboards or data teams only? Etc.

Search and Discovery: The first level of any data governance strategy is making sure relevant people can find the relevant datasets to do their analysis or build their AI model. If you don't implement this step, companies end up with a lot of questions on Slack and useless meetings with the engineering teams. The company ends up with a lot of duplicate tables, analyses and dashboards. It takes valuable time to engineering resources that are needed to perform the next steps.

Metadata and Documentation: Once you can efficiently find the data. You need to understand it quickly in order to assess if it is going to be useful. For example, you are looking at a dataset called "active_users_revenue_2021". There is a column "payment". Is this column in € or $? Has it been refreshed this

morning, last week, or last year? Does it contain all the data on active users or just the ones in Europe? If I remove a column, will this break important dashboards for the marketing or finance team? Etc.

Data Quality: Now that you have data, stored in scalable infrastructure, that everyone can find and understand, you need to trust that what is inside is of high quality. This is why so many data observability and reliability tools were born in the last five years. Data observability is the general concept of using automated monitoring, alerting, and triaging to eliminate data downtime. The two main approaches to data quality are: declarative (manually define thresholds and behaviour) or ML- driven (detecting sudden changes in distribution).

Security and Access right: Some data might be more private or strategic than others: you need to improve security as well as possible. Let's say you are a bank, you don't want to give access to the transaction logs to anyone in the company. You need to define access rights and managing them efficiently can quickly become a struggle as the number and type of people working with data grow. Sometimes, you want to give access to someone for a specific mission and for anything else. What happens when one of your employees was in the finance department but moves to marketing? You need a program to manage these rights thoroughly and efficiently to ensure key security standards.

Compliance and Regulation: This one is self-explanatory. To comply with various policies and regulations, you need to list all assets, report on personal information and usage to comply with regulation. For now, only enterprise companies are targeted by regulators, it is just a question of time before smaller companies start receiving fines. In most organizations, a yearly committee helps to drive the governance program.

IN ADDITION, INSTITUTIONAL OVERSIGHT

Data protection and privacy in general, and with respect to ID systems, are often subject to the oversight of an independent supervisory or regulatory authority to ensure compliance with privacy and data protection law, including protecting individuals' rights. The supervisory authority might be a single government official, ombudsman or a body with several members.

Genuine independence of such an authority is a key factor, with independence being measured by structural factors such as the composition of the authority, the method of appointment of members, the power and timeframe for exercising oversight functions, the allocation of sufficient resources and the ability to make meaningful decisions without external interference.

The supervisory authority may handle public complaints, even though every individual whose data is collected may have recourse to an external binding legal process and ultimately the courts at least on matters of law. In terms of remedies, the authority may have the power to oblige the ID system to rectify, delete or destroy inaccurate or illegally collected data.

Specifically, the Council of Europe (CoE) Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108, CoE 2018)—which was recently updated as Convention 108+—indicates that the powers and duties of such an authority may include:

1. duties to monitor, investigate and enforce compliance with individual privacy and data protection rights.
2. duties to monitor developments and their impact on individual privacy and data protection rights.
3. powers to receive complaints and conduct investigations of potential violations of individual privacy and data protection
4. rights.
5. powers to issue decisions on violations of such rights and order remedial action or meaningful sanctions;
6. duties to promote public awareness of the rights of individuals and the responsibilities of those entities holding and processing personal data.

7. a duty to give specific attention to the data protection rights of children and other vulnerable individuals.
8. issuing opinions prior to the implementation of data processing operations.
9. advising on legislative or administrative measures.
10. recommending codes of conduct or referring cases to national parliaments or other state institutions.
11. issuing regular reports, publishing opinions and other public communications to keep the public informed about their rights and obligations and about data protection issues in general.

DATA SHARING

Because the linkage of information across databases intensifies privacy and data protection concerns, legal frameworks can mitigate risks by stipulating all the purposes for which personal data in an ID system is shared, by both government and non-government entities. In addition, public entities may be limited to obtaining specific information justified by their functions (i.e., the "need-to-know" principle).

BENEFITS OF SHARING INFORMATION

1. convenience for both government and citizen
2. better government service delivery.
3. improved efficiency through more effective use of data.
4. seamless service transfer when data subjects change address.
5. improved risk management.
6. cost savings as duplication of effort is eliminated.

However, information-sharing between government agencies, if not well-regulated, can turn into a "back door" which allows circumvention of individual privacy and data protection safeguards. Comprehensive population databases, like those established as part of ID systems, are a tempting resource for law enforcement authorities, particularly when they contain biometrics.

Particular concerns arise in relation to collection of DNA information which, like other biometric data, may be used not only for the purposes of identifying an individual, but also as evidence in the process of investigating whether he or she has committed a crime.

This type of information sharing can take place even without the technological compatibility of interoperability. For example, police could contact ID officials and ask them to pull the record of a particular individual and share information such as fingerprints, facial image, address or names of family members.

The security of personal data transferred across national borders has been one of the drivers for international consensus on the fundamental principles for the protection of personal data.