REPUBLIC OF KENYA

OFFICE OF THE DATA PROTECTION COMMISSIONER

Guidance Notes on historical & Statistical Purposes

2025

TABLE OF CONTENTS

FOREWORD

The protection of personal data is a constitutional right under Article 31 of the Constitution of Kenya, 2010, and a cornerstone of the Data Protection Act, 2019. As the custodian of this right, the Office of the Data Protection Commissioner (Office) is mandated to regulate the processing of personal data to ensure that it is undertaken lawfully, fairly, and transparently.

Historical and statistical processing of personal data plays a vital role in preserving the nation's heritage, informing public policy, and advancing research. Archival records, cultural documentation, and official statistics contribute to collective knowledge, national development, and evidence-based decision-making. However, such processing also presents unique privacy challenges, including the risk of re-identification, prolonged retention of personal data, and the historical handling of sensitive information.

This Guidance Note therefore provides clarity to data controllers and processors on their obligations when processing personal data for historical and statistical purposes. It outlines the applicable lawful bases under the Data Protection Act, 2019, including consent, public interest, legal obligation, and functions of a public authority. It further underscores the safeguards necessary to protect personal data from unauthorized access, data breaches, and misuse.

The Office affirms that exemptions under the law for historical and statistical purposes are not absolute, but conditional. They must be exercised with proportionality, necessity, and demonstrable public interest, while upholding the principles of data protection. By adopting the practices set out in this Guidance Note, entities will not only comply with the law but also strengthen public trust, protect the dignity of individuals, and ensure that Kenya's historical and statistical record-keeping meets both national and international standards.

It is my expectation that this Guidance Note will serve as a practical reference for public and private institutions, researchers, and archivists, enabling them to balance the advancement of knowledge and preservation of national heritage with the fundamental rights of data subjects. In this way, Kenya can preserve its past, inform its present, and guide its future in a manner that is both responsible and privacy-respecting.

Immaculate Kassait, MBS
Data Commissioner

DEFINITION OF TERMS

"Act" means the Data Protection Act, No 24. of 2019.

"Anonymisation" means the removal of personal identifiers from personal data so that the data subject is no longer identifiable.

"Data Commissioner" means the person appointed pursuant to section 6 of the Act.

"Data Controller" means a natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purpose and means of Processing of Personal Data;

"Data Processor" means a natural or legal person, public authority, agency or other body which processes Personal Data on behalf of the Data Controller.

"Data Handler" means a data controller and or data processor.

"Data Subject" means an identified or identifiable natural person who is the subject of Personal Data.

"Entity" or **"Entities"** means a natural (individual) or legal person, public authority, agency or other body that processes (handles) Personal Data.

"Establishment documents" includes

(a) a Statute, Charter or statutory instrument in which a body is established;

(b) registration certificate;

(c) trust deeds in which a trust has been established; and

(d) other instruments by which a body is established including its governing and administrative structure.

For purposes of this guidance note, the Establishment document required is a list of all entities under the registering entity.

"Office" means the Office of the Data Protection Commissioner as established in section 5 of the Act.

"Personal Data" means any information relating to an identified or identifiable natural person.

"Processing" means any operation or sets of operations which is performed on Personal Data or on sets of Personal Data whether by automated means or not, such as:

(a) collection, recording, organization, structuring;

(b) storage, adaptation or alteration;

(c) retrieval, consultation or use;

(d) disclosure by transmission, dissemination, or otherwise making available; or

(e) alignment or combination, restriction, erasure or destruction

"Pseudonymisation" means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, and such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data is not attributed to an identified or identifiable natural person.

"Regulations" means the Data Protection General Regulations, 2021 and all other regulations enacted in accordance with section 71 of the Data Protection Act

"Sensitive Personal Data" means data revealing the natural person's race, health status, ethnic social origin, conscience, belief, genetic data, biometric data, property details, marital status, family details including names of the person's children, parents, spouse or spouses, sex or the sexual orientation of the Data Subject.

OFFICE

The Office of the Data Protection Commissioner (Office) is a government agency established to protect the privacy and security of personal data in our increasingly digital world. It is responsible for enforcing data protection laws and policies to safeguard the privacy, dignity, and fundamental rights of individuals. The Office is mandated to oversee the implementation and enforcement of the Data Protection Act, 2019, which regulates the processing of personal data of persons located in Kenya by both private and public sector organizations.

The Office plays a vital role in ensuring that individuals have control over their personal data and that organizations respect their privacy rights. The office's work involves monitoring and enforcing compliance with data protection regulations, investigating data breaches, and imposing sanctions on entities that violate data protection laws. In addition, the Office is responsible for raising public awareness about data protection issues and educating individuals and organizations on how to protect personal data. With the growing importance of data protection in our digital age, the Office of the data protection commissioner is a critical institution in maintaining trust and confidence in our data-driven society.

The Office of the Data Commissioner is uniquely positioned to facilitate both the government and private sector entities in achieving Government's strategic goals under the "Bottom-Up Economic Transformation Agenda" and, in particular, its digital superhighway initiative. As the digital landscape expands, the need for robust data protection mechanisms becomes paramount. The Office, with its mandate to oversee, regulate, and ensure lawful data processing, plays a pivotal role in this transformation. Kenya remains at the cutting edge of digital transformation while maintaining stringent data protection standards. The Office of the Data Commissioner serves as a key stakeholder and regulator in guiding the nation's digital superhighway journey by ensuring that as we advance technologically, the rights and privacy of individuals remain safeguarded.

INTRODUCTION

The purpose of this Guidance Note is to provide clarity and direction on the processing of personal data for historical and statistical purposes. Given the increasing importance of data in research and public information, it is essential to balance the need for processing data with the protection of individuals' privacy rights. This guidance aims to ensure compliance with the Data Protection Act, 2019, and to promote best practices among data controllers and processors.

The notion of statistical purposes refers to any data processing operation necessary for statistical surveys or to produce statistical results. These results may be expressed numerically, such as the percentage of women within some organisation or may be used in non-numerical form. The data is presented as aggregated data that do not consist of personal data, that is data that may be used to identify a particular individual. Collection of data for statistical purposes, as well as the use thereof, should be done in accordance with best practices. Statistical purposes can be pursued either by the public or the private sector. However, there is a difference between official statistics, which are often subject to stringent rules, and other statistical purposes. It is important to note that in certain cases, the personal data processed in order to reach statistical results may be kept for a certain period to prove that the process was performed in accordance with set rules.

The notion of historical purposes refers to any data processing operation necessary for the documentation, preservation, and study of past events, activities, or experiences. This type of processing is crucial for understanding historical contexts, analysing past trends, and preserving cultural heritage. Unlike statistical purposes, which often result in numerical data, historical purposes typically involve qualitative data, including narratives, documents and images.

Collection of data for historical purposes should adhere to best practices, ensuring accuracy, authenticity, and ethical standards in documentation and preservation. Historical research can be conducted by public institutions, such as archives, museums, and universities, as well as private entities, including historians and researchers.

It is also important to distinguish between official historical records, which are often subject to stringent archival rules, and other historical purposes. Official historical records may include government documents, official correspondences, and public records, which require careful handling and long-term preservation.
In some cases, personal data processed for historical purposes may need to be retained for extended periods to maintain the integrity of historical records and to ensure that the data can be revisited and reanalyzed in the future. This retention must comply with legal requirements and ethical standards, ensuring that the data is protected and used responsibly

PRIVACY CONCERNS

Processing personal data for historical and statistical purposes presents unique privacy challenges. These concerns must be addressed to respect individuals' privacy rights while enabling valuable research and data analysis:

One significant challenge is the risk of re-identification from ostensibly anonymized data. Even when direct personal identifiers are removed, it may still be possible to re-identify individuals by combining the data with other data sets, especially those containing detailed or unique information. This risk is particularly high with modern re-identification techniques.

Historical and statistical data often contain sensitive information, such as health records, financial details, or personal beliefs. Processing such data can have serious implications for the privacy and dignity of individuals, especially if the data pertains to vulnerable populations or stigmatizing conditions.

Ensuring the security of historical and statistical data is crucial to prevent unauthorized access, data breaches, and misuse. Data controllers and processors must ensure secure data-sharing protocols and implement robust security measures, including encryption, access controls, and regular security audits. Failure to secure data adequately can result in significant privacy breaches.

SCOPE AND PURPOSE OF GUIDANCE NOTE

The Office of the Data Protection Commissioner is the regulatory office established pursuant to the Data Protection Act, 2019 ("the Act") and charged with the responsibility of exercising oversight on data processing operations to ensure that the processing of personal data of data subjects located in Kenya is carried out in accordance with the Act.

The scope of this guidance note is to outline the obligations of data controllers and processors regarding the processing of personal data for historical and statistical purposes. It provides:

a. An overview of the lawful basis for processing such data.
b. Guidance on further processing where data was not initially collected for historical or statistical purposes.
c. Information on exemptions under sections 51, 52, and 53 of the Data Protection Act.
d. Implementation of data protection principles.
e. Best practices for data processing.

This Guidance Note considers:

a. The Data Protection Act, 2019;
b. The Data Protection Regulations;
c. The Data Protection Policy, 2018;
d. Other Applicable laws and regulations and
e. International Best Practice.

LEGISLATIVE FRAMEWORK

Public Archives and Documentation Services Act
This section details the responsibilities for archiving public documents and the role of the National Archives in preserving and providing access to these documents. This legal framework ensures that all significant governmental and public sector documents are preserved for posterity and are accessible for historical and research purposes.

Statistics Act, 2006
Provides a comprehensive legal framework for the collection, analysis, and dissemination of statistical data in Kenya. The Act outlines the roles, responsibilities, and powers of the Kenya National Bureau of Statistics (KNBS), the primary agency responsible for managing statistical data in the country.

LAWFUL BASIS FOR PROCESSING PERSONAL DATA FOR HISTORICAL AND STATISTICAL PURPOSES.

Section 30 of the Act, outlines several lawful bases for processing personal data, which include:

7.1 Consent

Processing personal data is lawful when the data subject gives explicit, informed, and freely given consent for one or more specified purposes. This ensures that individuals are aware of and agree to how their data will be used. However, while consent is a valid lawful basis, its suitability for long-term historical or statistical purposes — especially where future research objectives are not fully defined — should be carefully evaluated. In such cases, alternative lawful bases such as public interest or legal obligation may offer a more sustainable and appropriate foundation for processing.

Example: A historian plans to conduct a historical study on local migration patterns. They seek consent from individuals to use their personal letters, photographs, and stories, providing detailed information about the research purpose and how the data will be used. Participants sign consent forms, ensuring they are fully informed and agree to the processing of their personal data.

7.2 Legal Obligation

Processing personal data is lawful if it is necessary for compliance with a legal obligation to which the data controller is subject. This basis applies when an entity is required by law to process certain data, ensuring adherence to statutory duties and regulations. It provides a clear framework under which data processing is mandated by legal requirements.

Example:
Government departments are required to submit certain records to the National Archives under the Public Archives and Documentation Services Act. The National Archives processes these documents, which may contain personal data, to comply with the legal obligation of preserving government records. This ensures important historical documents are maintained according to statutory requirements.

7.3 Public Interest

The Processing personal data is lawful if it is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller. This basis is often used by public bodies and organizations that perform tasks benefiting society, such as conducting censuses or public health research.

Example:
A national statistics office conducts a population census to gather comprehensive demographic data. This task, carried out in the public interest, provides essential information for public policy and resource allocation. The data collected helps government bodies make informed decisions that benefit the entire population.

7.4 Performance of a Task by a Public Authority

Processing personal data is lawful if it is necessary for the performance of any task carried out by a public authority. This basis is specific to activities undertaken by governmental and public sector bodies as part of their official functions, ensuring they can perform their statutory duties effectively.

Example:
Local government processes residential data to analyze housing needs and allocate resources appropriately. This data collection and analysis are part of the local government's duty/mandate to ensure adequate housing for its residents. By processing this data, the local government can identify areas with housing shortages and develop policies to address these needs.

7.5 Functions of a Public Nature

Processing personal data is necessary for the exercise, by any person in the public interest, of any other functions of a public nature. This basis applies to activities serving the public good, even if carried out by private entities, provided they align with the public interest.

Example:
A cultural heritage organization processes personal data to document and preserve local traditions and histories. This work is done in the public interest, ensuring that valuable cultural and historical knowledge is preserved for future generations. The organization collects personal stories and photographs making sure the data is used responsibly and for the intended public benefit.

7.5.1 Archiving

The Public Archives and Documentation Services Act mandates specific obligations for archiving documents in the public interest. This provision is crucial for ensuring the preservation and accessibility of public records for historical and research purposes. Public authorities and state corporations must comply with these legal requirements to supply copies of documents and reports to the National Archives.

The Public Archives and Documentation Services Act ensures that all significant governmental and public sector documents are preserved for posterity and are accessible for historical and research purposes. By adhering to these guidelines, public authorities and state corporations can maintain a comprehensive and accessible archive of public records, balancing the need for preservation with the protection of sensitive information. This approach not only supports historical research and public information initiatives but also ensures compliance with legal requirements and enhances the integrity and reliability of archival practices.

7.6 Legitimate Interests

Processing personal data is lawful if it is necessary for the legitimate interests pursued by the data controller or a third party, except where such interests are overridden by the rights and freedoms of the data subject. This basis requires a careful balancing test to ensure the processing is justified and does not unduly impact the data subject's rights.

Example:
A private archive processes personal data in historical documents to provide research services to scholars. The archive's legitimate interest in preserving and making historical data accessible balances against the need to protect individual privacy. The archive anonymizes sensitive information where possible and restricts access to authorized researchers to minimize any potential privacy impact.

7.7 Historical and Statistical Purposes

Processing personal data is lawful if it is necessary for historical and statistical purposes. This basis acknowledges the importance of such activities in advancing knowledge and understanding while requiring appropriate safeguards to protect personal data. It is important to note that when relying on Section 53 of the Act for historical and statistical purposes, the personal data must have been originally collected under a separate, primary lawful basis. This provision serves as a secondary lawful basis for the further processing of already collected personal data, exclusively for historical and statistical purposes.

Example:
A university uses student information that they already collected during admissions to develop statistical research that will inform financial and infrastructure strategy for future admissions. The data, originally collected for other purposes, is repurposed for statistical research, and the university ensures that the data is anonymized.

7.8 Further Processing

When personal data initially collected for a different purpose is to be used for historical or statistical purposes, it is essential to ensure that the further processing is compatible with the original purpose. This involves several key steps and safeguards to protect the rights of data subjects and maintain the integrity of the data processing activities.
First, it is crucial to assess the compatibility of the new processing purpose with the original intent for which the data was collected. This involves a thorough review to determine whether the new purpose aligns with the expectations of the data subjects and the initial context of data collection.

Transparency is fundamental when conducting further processing for historical, statistical, or research purposes. Data controllers should clearly inform data subjects, at the time of collection or before such further processing, about the intended secondary uses and the specific purposes involved. Even where consent is not required due to the compatibility of purposes or anonymization of data, transparency remains essential to uphold the principles of accountability and fairness.

Next, implementing appropriate safeguards is essential to protect the rights of data subjects. This includes measures such as anonymization to prevent the re-identification of individuals and robust access controls to limit data access to authorized personnel only. Anonymization ensures that identifiers are permanently removed or masked, significantly reducing the risk of privacy breaches. Access controls, on the other hand, help secure the data and ensure that only those with legitimate reasons can access it.

Finally, when publishing data for historical or statistical purposes, it is important to ensure that the data does not identify individuals. Data should be published in an aggregated form that does not reveal individual identities, thereby protecting the privacy of data subjects. If

identifiable data needs to be published, consent from the individuals must be obtained or another lawful basis has been established.

Example:

A historian is preparing a book on political developments over the past decade, focusing on the actions of former public officials. To support this work, the historian accesses official government archives and publicly available online sources. In the process, personal data related to both the officials and other individuals involved in the events is collected.

The historian takes care to ensure the data is used solely for historical research purposes and applies appropriate safeguards, such as anonymizing identities of private individuals where possible. Any information published in the final work is either already publicly available, anonymized, or disclosed onon a valid legal basis—balancing historical accuracy with respect to individual privacy.

OBLIGATIONS OF DATA CONTROLLERS FOR PROCESSING PERSONAL DATA FOR STATISTICAL AND HISTORICAL PURPOSES

Data handlers have several obligations to ensure compliance with the Act when processing personal data for statistical and historical purposes. These obligations are designed to protect the rights and privacy of data subjects while enabling the effective use of data for valuable research and analysis. Below are the key obligations of data handlers:

8.1. Lawful Basis for Processing

Data controllers must identify and document the lawful basis for processing personal data. Only one lawful basis should be relied upon for each processing activity, and this basis must be clearly communicated to data subjects. Data handlers should also inform data subjects of their right to withdraw consent and the consequences.

Example:

When relying on consent, data controllers must ensure that it is explicit, informed, and freely given. They should provide clear information about the purpose of the data processing and obtain documented consent from data subjects.

8.2. Data Minimization

Data controllers should collect only the personal data that is necessary for specified historical or statistical purposes. They should avoid collecting excessive or irrelevant data. Additionally, data handlers should document the necessity of each data item collected and review periodically for continued relevance.

8.3. Data Quality

Data controllers must ensure that the personal data they process is accurate, complete, and up to date. To maintain data quality, controllers should conduct regular reviews and must also enable data subjects to access their personal data and request corrections where necessary.

8.4. Transparency and Communication

Data controllers must inform data subjects about the purpose of data collection, the lawful basis for processing, and their rights. This information should be provided in a clear and accessible manner.

8.5. Data Security

Data controllers must implement appropriate technical and organizational measures to protect personal data from unauthorized access, loss, alteration, or destruction. This includes using encryption, access controls, regular security audits and employee training.

a. Anonymization and Pseudonymization

When processing personal data for historical and statistical purposes, data handlers should prioritize anonymization and pseudonymization techniques to protect the identity of data subjects. These methods significantly reduce the risk of re-identification and enhance privacy, ensuring compliance with data protection regulations and safeguarding individuals' rights.

i. Anonymization

Anonymization is the process of rendering personal data irreversibly unidentifiable, such that the data subject can no longer be identified, directly or indirectly, by any means reasonably likely to be used. True anonymization removes the data from the scope of data protection laws, as it no longer constitutes personal data. However, data handlers must be aware of the high bar for effective anonymization and the ongoing risk of re-identification, especially with evolving technologies.

Steps for Effective Anonymization:

i. **Identify Personal Data**: Determine which data elements could directly or indirectly identify individuals.
ii. **Remove Identifiers**: Remove or make obscure direct identifiers, such as names, addresses, and identification numbers.
iii. **Generalize Data**: Convert specific data points into broader categories (e.g., age ranges instead of exact ages).
iv. **Suppress Data**: Omit data points that could lead to identification, especially in small datasets.
v. **Aggregate Data**: Combine data into groups such as based on various categories or time periods to prevent the identification of individual records.
vi. **Test Anonymization**: Assess the data to ensure that individuals cannot be re-identified using reasonably available information.

Example
A university conducting a historical study on population health anonymizes patient records by removing names, addresses, and exact birthdates, and generalizes age into ranges (e.g., 20 - 30, 31 - 40). This ensures that the data can be used for research without risking the identification of individuals.

ii. Pseudonymization

Pseudonymization is the process of replacing private identifiers with fake identifiers or pseudonyms. Unlike anonymization, pseudonymized data can potentially be re-identified by using additional information kept separately. Pseudonymization enhances privacy while still allowing data to be linked back to the data subject, if necessary, under strict controls.

Steps for Effective Pseudonymization:

i. **Replace Identifiers**: Substitute identifiers with pseudonyms (e.g., replace names with unique codes).
ii. **Separate Data**: Store the pseudonymization key (the mapping between real identifiers and pseudonyms) separately and securely.
iii. **Access Controls**: Restrict access to the pseudonymization key to authorized personnel only.
iv. **Data Minimization**: Ensure that the pseudonymized dataset contains only the information necessary for the analysis.
v. **Regular Review**: Periodically review and update pseudonymization techniques to address evolving risks and challenges.

Example
An institution studying historical employment trends pseudonymizes employee records by replacing names and identification numbers with unique codes. The key linking these codes to real identities is stored separately and securely, accessible only to a limited number of authorized staff.

8.6. Purpose Limitation

Personal data should only be processed for the specific purposes for which it was collected. Any further processing for historical or statistical purposes must be compatible with the original purpose.

8.7. Storage Limitation

Data handlers must not retain personal data for longer than necessary. They should establish clear data retention periods based on the purposes of processing and securely delete or anonymize data once it is no longer required. For data intended for indefinite retention for historical purposes, this must be justified by a clear public interest and subject to robust anonymization or pseudonymization techniques, along with ongoing security measures.

8.8. Accountability and Compliance

Data handlers should maintain records of their data processing activities, including the lawful basis for processing, consent forms, data protection impact assessments, and any measures taken to protect data.
Further, data handlers should conduct regular audits and reviews of data processing activities to ensure ongoing compliance with data protection laws.

8.9. Rights of Data Subjects

Data controllers must facilitate the effective exercise of data subjects' rights by providing clear, accessible channels and establishing procedures to handle requests efficiently. These procedures should ensure that all requests are addressed within the statutory time frames set out in the Act.

Right	What does this mean?	Are there limits to this right?
Right to be informed	Right to be provided with extensive information including all their rights, contact details, source, retention period, purposes, categories and recipients etc.	This right is absolute; the individual whose data is being collected must be notified within reasonable timeframes, even if the data is obtained from secondary sources.
Right of access	Right to access personal data within 7 days and for free. It must be possible to make electronic subject access requests.	Data subjects have the right to access their personal data. However, certain data can be withheld if disclosure adversely affects the rights and freedoms of other data subjects.
Right to rectification (of inaccurate data)	A data subject may request correction of any personal information held to ensure that the data is accurate, true, or complete. Rectification of personal data should be within 14 days.	If a right to rectification is denied, the data handler should notify the subject of the refusal and provide reasons for the decline
Right to erasure "to be forgotten"	A data subject may Request for erasure if; • the data is no longer necessary for the purpose for which it was originally collected, • An individual withdraws their consent, • individual objects to the processing of their data,	Right does not apply if processing is necessary: • for archiving purposes in the public interest, • Scientific research, • necessary for freedom of expression or information, • public interest in public health, • fulfilment of legal claims,

	• the processing of data is for direct marketing and the individual objects, • the processing has been unlawful. Data controller or processor should respond within 14 days.	• Historical, research or statistical purposes where erasure is likely to render impossible or seriously impair the achievement of that processing.
Right to restrict processing	Right to request that processing be restricted where data cannot be deleted as it is required for legal reasons. This is a more limited right than the right to erasure. It allows controllers to quarantine data to be used solely for a limited range of purposes such as handling legal claims.	When declined the controller/processor shall communicate in writing to the person and provide reasons for objection.
Right to data portability	This allows individuals to move, copy, or transfer data easily from one controller/processor to another. The request should be acted on within thirty (30) days of receipt and upon payment of any reasonable charge.	The right to data portability only applies; to personal data that an individual has personally provided, where the processing is based on consent or the performance of a contract or where processing is carried by automated means (i.e. excluding paper files)
Right to object to processing	Right to object to the processing of all or part of their data; generally, or for a specified purpose, objection is absolute where processing is for direct marketing. A controller should respond within 14 days.	The right does not apply where there are legal claims requiring data use, and when data controllers demonstrate a compelling legitimate interest that outweighs an individual's privacy interests for continued processing of personal data.
Right to withdraw consent	Data subjects have a right to withdraw consent as easily as it was to give consent.	The right is not applicable when personal data is anonymized or pseudonymized, or when research analysis is completed, and conclusions are anonymized.
Right not to be subject to a decision based solely on automated	Data subjects have a right not to be evaluated or subject to decisions based solely on automation where decisions have legal or significant effects.	The right does not apply if the decision is based on explicit consent; necessary for a contract; authorised by law.

decision making		

8.10. Data Protection Impact Assessments (DPIAs)

For processing activities that are likely to result in a high risk to the rights and freedoms of data subjects, data controllers should conduct Data Protection Impact Assessments (DPIAs) to identify and mitigate risks.

A guidance note on how to conduct a DPIA is available on https://www.odpc.go.ke/

8.11. Data Breach Notification

In the event of a data breach, data controllers must notify the Data Commissioner within 72 hours and, where appropriate, the affected data subjects without undue delay. They should have tested and documented incident response plans in place to manage breaches effectively.

8.12. Collaboration with Data Processors or other Data Controllers

When engaging with data processors, data controllers must ensure that processors comply with data protection requirements. This includes establishing binding data processing agreements that outline the responsibilities and obligations of the processors, and a provision for regular audit by the controller to verify compliance. If routinely sharing data with other data controllers, a data sharing agreement will be necessary to safeguard data subject rights.

EXEMPTIONS

There are no absolute exemptions under the Act. Exemptions must be justified based on proportionality, necessity, and demonstratable public interest. Therefore, reliance on exemptions for historical or statistical purposes does not absolve a data controller or processor from complying with key data protection principles. These include:

a) Establishing a lawful basis for processing personal data;
b) Collecting only the data necessary and proportionate to the intended purpose;
c) Ensuring the accuracy and quality of the data;
d) Upholding ethical standards in data processing; and
e) Implementing appropriate security measures to protect personal data.

In addition, data controllers and processors must ensure that the processing of personal data for historical or statistical purposes is strictly limited to those purposes and that any publication or disclosure does not reveal data in an identifiable form.

9.1. Historical Purposes

Historical research often necessitates access to personal data contained in archives and historical records. Institutions providing access to such records must ensure that the research is legitimate. This typically involves verifying that the researcher is affiliated with a recognized academic or research institution and that the study has been approved through an appropriate ethics or research review process. Such validation helps ensure that the processing is in the public interest and that appropriate safeguards are in place.

Importantly, the exemption is not limited to professional historians alone—it may also apply to other individuals or entities conducting bona fide historical research, including independent researchers, journalists, or public bodies, provided they can demonstrate a legitimate purpose and adhere to applicable data protection safeguards.

Data controllers and processors must ensure that any data made available or published as part of such research is where applicable anonymized and does not identify individuals, unless explicit consent has been obtained. This framework balances the advancement of historical knowledge with the obligation to protect individual privacy.

For example, an organization conducting a study on early 20th-century public figures may use archival records, ensuring that any disclosures are made in a non-identifiable format unless consent has been explicitly secured.

9.2. Statistical Purposes

Statistical data processing involves collecting and analyzing data to produce numerical or non-numerical results that provide insights into various trends and patterns. This exemption permits

the processing of personal data if necessary for producing statistical results, provided the data is presented in aggregated form to avoid identifying individuals. Researchers must follow best practices and adhere to stringent rules, especially official statistics. For example, a government agency analyzing employment rates must aggregate the data to ensure individuals cannot be identified in the published results.

Section 53 of the Data Protection Act, 2019, provides specific exemptions for the processing of personal data for historical and statistical purposes. These exemptions recognize the unique value of such data processing activities and their potential public benefits. However, they also emphasize the need for appropriate safeguards to protect the rights and privacy of data subjects. Below is a detailed explanation of the Section 53 exemption and an assessment framework to determine the public interest in processing such data.

EXEMPTIONS FOR HISTORICAL AND STATISTICAL PURPOSES

10.1. Compatibility of Further Processing

Section 53(1) states that the further processing of personal data for historical or statistical purposes is considered compatible with the original purpose for which the data was collected. This means that data initially collected for one purpose can be further processed for historical or statistical analysis without obtaining new consent, provided the processing is necessary, and appropriate safeguards are implemented.

Data handlers must ensure that such further processing does not adversely affect the rights of data subjects. Measures such as data anonymization, minimizing data collection, and restricting access to authorized personnel are critical to protecting individuals' privacy.

Example.

A hospital originally collects patient data for medical treatment. This data is later used for historical research on disease trends over the past century. The hospital anonymizes the data, ensuring no individuals can be identified in the research findings.

10.2. Anonymization and Secure Data Handling

Section 53(2) requires that personal data processed for historical or statistical purposes be anonymized wherever possible. This minimizes the risk of re-identification and protects individuals' privacy.

Data Handlers must, therefore, implement industry-standard robust anonymization techniques, conduct periodic risk assessment of anonymization techniques, restrict data access to authorized personnel, and use secure data storage methods to prevent unauthorized access or data breaches.

Example.

A statistical agency anonymizes employment data before using it for a study on labor market trends. Only aggregated results are published, and access to the raw data is limited to a small team.

10.3. Non-Identifiable Data for Publication

Section 53(3) prohibits publishing personal data processed for historical, statistical, or research purposes in an identifiable form. Results must be anonymized to prevent identification of data subjects. If there is a need to use or publish identifiable data, this must be done outside the exemption, based on a separate lawful basis such as explicit consent or another condition under Section 30 of the act.

When publishing research findings, ensure that data is presented in an aggregated or anonymized form. Obtain explicit consent from individuals if identifiable data must be disclosed.

Example:

A paper on historical migration patterns uses aggregated data to show trends without revealing individual identities. If the study needs to include specific personal stories, the researchers obtain explicit consent from those individuals or their descendants.

10.4 Public Interest Assessment Framework

Determining whether the processing of personal data for historical or statistical purposes is in public interest involves a thorough assessment. This framework can guide data controllers and processors in making that determination:

a) Define the Purpose and Benefits

Clearly articulate the purpose of the data processing activity and the expected benefits to society. This includes understanding how the research or analysis will contribute to public knowledge, policy-making, or cultural preservation.

Example:

An exercise aims to analyze the impact of urbanization on community health over the past century. The expected benefit is to inform current urban planning and public health policies.

b) Evaluate the Necessity and Proportionality

Assess whether the processing is necessary to achieve the stated purpose and whether it is proportionate to the potential impact on data subjects. Consider if there are less intrusive means to obtain the same results.

Example:

The study requires historical health records to analyze trends accurately. The data controller evaluates that anonymizing the data and restricting access are sufficient measures to ensure the processing is proportionate and necessary.

c) Implement Safeguards and Mitigations

Data handlers must identify and implement appropriate safeguards to protect the rights and privacy of data subjects when processing personal data for historical, statistical, or research purposes. These safeguards may include data anonymization or pseudonymization, strict access controls, data minimization, secure storage, and clear protocols for data handling and retention.

Example:

A data controller conducting historical research on early public health initiatives anonymizes personal identifiers in archived health records before making them available to researchers. Access is restricted to a vetted group of researchers affiliated with recognized academic institutions. The data is encrypted, stored in a secure environment, and subject to regular audits to ensure ongoing compliance with data protection requirements and ethical research standards.

d) Consider public expectations and transparency

Take into account the reasonable expectations of data subjects regarding the use of their data. Ensure transparency by informing data subjects (or their representatives) about the processing activities and the measures in place to protect their data.

Example:

The organisation informs the community about the study through public notices and meetings, explaining how the data will be used and the protections in place. This transparency helps build trust and manage expectations.

e) Balance Public Interest Against Potential Harm

Conduct a balancing test to weigh the public interest in the data processing against any potential harm to data subjects. This includes considering the likelihood and severity of any adverse effects on individuals.

Example.

The potential public benefit of understanding urbanization's impact on health is significant, and the risks are mitigated through anonymization and secure data handling. The data controller concludes that the public interest outweighs the potential harm.

f) Document the Assessment

Maintain comprehensive records of the public interest assessment, including the rationale for the processing, the safeguards implemented, and the outcomes of the balancing test. This documentation is essential for accountability and compliance with data protection regulations.

Example:

The data controller documents the entire assessment process, including the purpose of the study, the benefits, the safeguards, and the justification for determining the processing is in the public interest.

ANNEX A- CHECKLIST FOR COMPLIANCE

No.	Compliance Requirement	Questions	Yes	No	Comments/ Remedial Actions
1.	Lawful basis	We have reviewed the purposes of our processing activities and selected the most appropriate lawful basis (or bases) for each activity.			
		We have checked that the processing is necessary for the relevant purpose and are satisfied that there is no other reasonable and less intrusive way to achieve that purpose.			
		We have documented our decision on which lawful basis applies to help us demonstrate compliance.			
		We have included information about both the purposes of the processing and the lawful basis for the processing in our privacy notice.			
		Where we process sensitive data, we have also complied with section 44, 45 and 46 of the Act.			
	Compliance obligation	Have you registered with the Office?			
		Have you complied with the principles of data protection as per section 25 of the Act?			
		Have you implemented data protection by design or by default?			
		Do you have any mechanism to implement data subject rights?			
		Have you implemented appropriate safeguards for the protection of the data?			
	Exemptions	Have you established whether your intended processing for historical or statistical purposes is compatible			

		with the original purpose for collecting the data?			
		Had you informed the data subjects at the original collection that their personal information may be used for further historical and statistical purposes?			
		Do you intend to publish anonymized or non-identifiable data?			
		Have you put in place adequate safeguards to protect the information?			