OFFICE OF THE DATA PROTECTION COMMISSIONER

Guidance Notes - Processing for Research Purpose

2025

TABLE OF CONTENTS

Foreword **4**

Definitions **5**

The Office **8**

Introduction **9**

Privacy Concerns **10**

Scope and Purposes of the Guidance Notes **13**

Legislative Framework **14**

Principles of Data Protection **16**

6.1 Lawfulness, Fairness & Transparency *16*

6.2 Purpose Limitation *16*

6.3 Data Minimization *17*

6.4 Accuracy *17*

6.5 Confidentiality & Integrity *18*

6.6 Storage Limitation *19*

6.7 Accountability *20*

Lawful basis for Processing Data for Research Purposes **22**

7.1 Consent *22*

7.2 Performance of a Task Carried out by a Public Authority *23*

7.3 Legitimate Interest *24*

7.3.1 Legitimate Interest Assessment Framework 24

7.4 Public Interest *26*

7.5 Scientific Research Purposes *27*

Processing of Sensitive Personal data **28**

8.1 Clinical Trials *28*

8.2 Assessing Vulnerability of Research Participants *28*

9. Compliance Obligations **31**

10. Rights of Data Subjects **33**

10. Exemptions **35**

10.1 Conditions for Exemptions *35*

10.2 Exemptions under Public Interest *35*

10.2.1 Public Interest Assessment Framework 36

10.3 Exemptions on further processing for Compatible Purposes 37

11. Secure Data Handling while Conducting Research 39

11.1 Anonymisation and Pseudonymisation 39
11.1.1. Anonymisation 39
11.1.2. Pseudonymisation 40

11.2. Non-Identifiable Data for Publication of Research Output 41

Appendix A: Compliance Checklist for internal use by Researchers. 42

FOREWORD

The right to privacy, enshrined under Article 31 of the Constitution of Kenya, 2010, is a fundamental right that must be protected in all domains of personal data processing. The Data Protection Act, 2019, operationalizes this right by providing a comprehensive legal framework to ensure that personal data is processed lawfully, fairly, and transparently. Within this framework, the processing of personal data for research purposes occupies a unique position—supporting the generation of new knowledge, innovation, and evidence-based policymaking, while raising important considerations for the safeguarding the right to privacy.

Scientific, academic, medical, technological, or social research relies and involves the processing of personal. Yet, if not properly governed, such processing may expose research participants to privacy violations, misuse of sensitive data or other harms, particularly when involving vulnerable groups. It is therefore essential that research institutions, universities, government agencies, non-governmental organisations, and private entities engaged in research adhere to both the letterand spirit of the Data Protection Act.

This Guidance Note has been developed to provide guidance to researchers and to support compliance with the Act and the Regulations. The Guidance Note provides clarity on the lawful bases for processing personal data for research purposes, outlines key safeguards that researchers should implement to protect personal data while conducting research and it highlights the obligations of researchers in respecting the rights of data subjects while recognizing that the specific exemption under research that may apply under defined circumstances. The ODPC reiterates that exemptions for research purposes are not blanket permissions, but conditional measures requiring proportionality, necessity, and demonstrable public benefit.

Compliance with this Guidance Note will enable researchers to balance the advancement of knowledge with the protection of individual dignity and rights. In so doing, Kenya will continue to nurture an environment of trust that supports innovation, scientific inquiry, and social progress.

It is my expectation that this Guidance Note will serve as an authoritative reference for all actors engaged in research, guiding them to process personal data responsibly, ethically, and lawfully, while upholding the constitutional and statutory rights of every Kenyan.

Immaculate Kassait, MBS
Data Commissioner

DEFINITIONS

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

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

"Child" means a child as defined in the Children Act Cap 141.

"Clinical Trial" means any research study involving human subjects undergoing biomedical, behavioural, or epidemiological activities, including data collection or analysis. Subjects may undergo manipulation, intervention, observation, or interaction with researchers, directly or through environmental changes, to evaluate the activities' impact and generate new knowledge.

"Compatible Purposes" means further processing of personal data that aligns with the original purpose for which the data was collected, such as research purposes under Section 53 of the Data Protection Act, 2019, provided appropriate safeguards are in place.

"Confidentiality" is the degree to which access and disclosure of given information is limited to authorized entities (e.g., users) and for intended purposes only thereby preventing access by or disclosure to unauthorized entities (users).

"Data" means information which -

a) is processed by means of equipment operating automatically in response to instructions given for that purpose;
b) is recorded intending to be processed by such equipment.
c) is recorded as part of a relevant filing system.
d) where it does not fall under paragraphs (a) (b) or (c), forms part of an accessible record; or
e) is recorded information which is held by a public entity and does not fall within any of paragraphs (a) to (d).

"Data Sharing Agreements" means written contracts between a data controller or processor and a third party that outline the terms, conditions, and safeguards for sharing personal data, ensuring compliance with the Data Protection Act, 2019.

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

"Informed consent" means any manifestation of express, unequivocal, free, specific and informed indication of the data subject's wishes by a statement or by a clear affirmative action, signifying agreement to the processing of personal data relating to the data subject.

"Integrity" is the degree to which access and disclosure of given information is limited to authorized entities (e.g., users) and for intended purposes only thereby preventing access by or disclosure to unauthorized entities (users).

"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.
e) alignment or combination, restriction, erasure or destruction.

"Public interest" concerns matter which may reasonably be judged to have an impact on the population, or groups of citizens. It includes but is not limited to protecting public health or safety; or preventing the public from being misled by an action or statement of an individual or organization or any other situation that may be determined from time to time.

"Pseudonymization" 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 organizational 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; the Data Protection (Complaints Handling and Enforcement Procedures) Regulations, 2021; and the Data Protection (Registration of Data Controllers and Data Processors) Regulations, 2021.

"Research" means a systematic investigation, including experimental development, undertaken on a purposeful basis and directed towards the generation of new knowledge. It includes academic, marketing, scientific, technological, clinical trials, product testing and development, polling activities and medical research.

"Research purposes" refers to the collection, analysis, and interpretation of personal data within a systematic investigation aimed at generating new knowledge.

"Researcher" means any natural or legal person who collects and carries out methodological processing of data with the objective of generating new knowledge. This may include but not

limited to research institutions, universities and other academic institutions, polling organizations, students, hospitals, policy organizations, licensed persons under the Science, Technology and Innovation Act, foreign scientists and entities carrying out research activities.

"**Scientific Research**" means research conducted for the purpose of contributing towards science, technology and innovation by the systematic collection, interpretation and evaluation of data.

"**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.

"**Vulnerable research participant**" means a person who needs special care, support or protection, including the orphaned and vulnerable children, widows or widowers, person with disability, elderly persons economically or educationally disadvantaged individuals, people in institutional settings (e.g., prisoners or hospital patients), cultural or ethnic minorities among other groups or indigent due to a risk of abuse or neglect. They are those persons who are relatively (or absolutely) incapable of protecting their own interests. More formally, they may have insufficient power, physical capacity, intelligence, education, resources, strength, or other needed attributes to protect their own interests.

THE OFFICE

The Office of the Data Protection Commissioner is a government agency established to protect the privacy and security of personal data in our increasingly digital world. The Office is mandated to oversee the implementation and enforcement of the Act, which regulates the processing of personal data, both by private and public organisations processing personal data of data subjects located in Kenya.

The Office plays a vital role in ensuring that individuals have control over their personal data and that organisations 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 organisations on how to protect personal data. With the growing importance of data protection in our digital age, the Office is a critical institution in maintaining trust and confidence in our data-driven society.

The Office 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 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 serves as a key stakeholder and regulator in guiding the nation's digital journey by ensuring that as we advance technologically, the rights and privacy of individuals remain safeguarded.

INTRODUCTION

The Office has developed this guidance note for processing personal data for research purposes to support compliance with the Act and the Regulations. The guidance note outlines the compliance requirements and provides information on the interpretation and implementation of the Act and the regulations as it relates to processing personal data for research purposes. By following these guidelines, researchers can ensure that their research activities are conducted in a responsible and ethical manner, maintaining the trust of participants and stakeholders while contributing to the advancement of knowledge.

The Act provides some exemption to application of the Act, including the processing only for research purposes as outlined under section 53 (3) of the Act. However, not all processing of data for research purposes automatically qualifies for exemptions. Such exemptions apply on a case-by-case basis. Exemptions should be determined on the need to balance the right to privacy with the right to freely express one's thoughts and ideas, i.e., freedom of expression. Thus, the exemption provided in the Act is not a blanket exemption.

This Guidance notes aims to balance the right to privacy under Article 31 of the Constitution of Kenya with the need to promote responsible research that advances knowledge, innovation, and evidence-based policy making."

This note applies to academic institutions, public and private research organizations, government agencies, and any data controller or processor engaged in scientific, historical, or statistical research.

PRIVACY CONCERNS

Processing personal data for research purposes raises significant privacy concerns, including the potential misuse of personal data, lack of transparency around data collection and processing of personal data. Research relies on processing personal data without which advancing knowledge and societal well-being would be hampered. Processing personal data for research purposes must be guided by a risk-based approach, explicitly outlining mitigation strategies for identified privacy threats to strike a balance between utilizing personal data for research purposes and safeguarding individual privacy..

Here are some of the key privacy concerns in this sector: -

a) **Transparency**: Failure to adequately explain the research purpose, data collection methods, or potential privacy risks to participants makes it difficult for individuals to participate in the research and provide informed consent. Researchers must be specific and transparent about the personal data they collect, the intended research purpose, and the potential privacy implications. This includes obtaining informed consent from data subjects, whenever applicable, and clearly outlining how the data will be anonymized or pseudonymized to minimize identifiability. Unless data subjects give informed consent about sharing of their personal data, pseudonymization is the default approach before sharing any form of data. Anonymization is mandatory when sharing data publicly for open access, or for secondary analysis.

b) **Sensitive Personal Data**: Research involving data, such as health information or data revealing race or ethnicity which when mishandled poses significant risk to the rights of data subjects being infringed.

c) **Vulnerable Groups:** Privacy concerns arise when reporting on vulnerable groups or those having special needs. Vulnerable groups should be given greater protection due to their limited or lack of capacity to give informed consent. Researchers are required to request consent from the data subjects and demonstrate a compelling justification for processing such data and implement robust safeguards to protect it. Additionally, when research involves vulnerable populations, such as low-income individuals, children or people with disabilities, researchers must be extra cautious to avoid exacerbating existing inequalities or causing further harm in compliance with principles of data protection.

d) **Data Minimization and Retention:** Collection of excessive data that may be unnecessary for the research purpose and/or retained indefinitely exposes participants to unnecessary risks of misuse of their data. Participants may feel their privacy is being

disregarded if they believe researchers are collecting data beyond the stated research purpose. This can erode trust in the research process and discourage participation. Researchers must collect data relevant to the research objectives. Clear data retention policies and procedures for secure disposal are required to mitigate privacy risks.

e) **Data Security:** Risk of re-identification and disclosure of personal data to unauthorised individuals or institutions raises privacy concerns. This could happen through poor anonymity of personal data, unencrypted communication channels, human error, or insufficient data sharing agreements. Weak passwords, inadequate access controls, and unencrypted data storage increases this risk. Malicious actors or simple technical errors can lead to data manipulation, compromising its integrity and potentially leading to inaccurate research findings. Inaccurate data can also harm participants if used for decision-making. Researchers must implement appropriate technical and organizational safeguards commensurate with the sensitivity of the data collected.

f) **Accuracy:** Errors, both intentional and unintentional, can compromise data validity during research process. These errors can occur throughout the process, from design flaws and participant selection to data entry and analysis. Bias can skew findings towards one group compared to another. Researchers must be cautious to prevent these errors, ensuring data protection principles are upheld.

g) General Misuse of Data: Refers to the inappropriate or unethical handling of personal data, which can undermine individuals' privacy rights. When researchers collect personal information, there is a risk that it may be used beyond the intended scope, such as for unauthorized analysis, sharing with third parties, or profiling without consent. This misuse can lead to breaches of confidentiality, discrimination, and loss of trust in research institutions. Safeguarding personal data through strict adherence to ethical guidelines and data protection regulations is crucial to mitigate these privacy concerns and ensure that research serves its intended beneficial purpose without compromising individual rights.

h) **Use of Emerging Technologies**: Use of new technologies such as Artificial Intelligence (AI) raises ethical concerns as well as privacy concerns. AI systems can be used to collect, analyse, or make decisions based on personal data. AI systems can perpetuate or amplify biases present in training data, leading to unfair or discriminatory outcomes. Researchers utilizing AI tools must ensure transparency about the use of such technologies, conduct thorough impact assessments, implement robust

safeguards to prevent unintended privacy violations, ensure that when AI is used to make automated decision on the data subjects, they should be informed and there should be human intervention in the decision making in order to protect the rights and privacy of the data subjects.

i) **Data Sharing and Transfers:** Sharing and transferring personal data between researchers, institutions, or across borders introduces risks related to loss of control over the data, unauthorized access, and potential breaches of confidentiality. Inadequate safeguards during data sharing or transfer can result in personal data being exposed to unintended recipients or jurisdictions with weaker privacy protections. Researchers must ensure that data sharing agreements and/or data processing agreements are in place to ensure that data recipients are compliant with the applicable data protection laws, and appropriate technical and organizational measures, such as encryption and access controls, and implemented to protect data throughout its lifecycle.

j) **Data Errors in Research:** Errors, both intentional and unintentional, can compromise data validity during research process. These errors can occur throughout the process, from design flaws and participant selection to data entry and analysis. Bias can skew findings towards one group compared to another. Researchers must be cautious to prevent these errors, ensuring data protection principles are upheld. Research institutions should adopt data governance best practices across the entire data Life cycle (From Inception, collection, processing, analysis, storage to archival) with clear Standard Operating Procedures (SOPs) at every stage.

SCOPE AND PURPOSES OF THE GUIDANCE NOTES

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 is carried out in accordance with the Act.

In exercise of the powers of the Data Commissioner pursuant to the Act and the Regulations, this Guidance Note provides for considerations that must be present when processing personal data for research purposes.

This guidance note considers:

a) The Data Protection Act, 2019;

b) The Data Protection Regulations, 2021;

c) Other applicable laws and regulations; and

d) International Best Practice.

This Guidance Note applies to Research Institutions, Researchers (academic and non-academic), Research assistants and data analysts, and data controllers and processors processing personal data for research purposes in either the public or private sectors and non-governmental organisations.

LEGISLATIVE FRAMEWORK

The processing of personal data for research purposes in Kenya is governed by:

The Constitution of Kenya 2010

a) **Article 31 (c) and (d)** guarantee the right to privacy regarding information relating to their family or private affairs unnecessarily required or revealed, or that of their communication infringed**.**

b) **Article 33** guarantees freedom of expression which includes academic freedom and freedom of scientific research

c) **Article 186** provides that oversight of research institutions is a function of the National Government.

Data Protection Act, 2019 - The Act provides for lawful basis of processing personal data for scientific research activities. It also provides exemptions for further processing of personal data for research purposes and the conditions for exemption.

Science, Technology and Innovation Act establishes the National Commission of Science, Technology and Innovation (NACOSTI) to oversee the integrity and standards of research conducted in Kenya. The Commission also has the power to issue licenses and approvals to researchers ensuring compliance with ethical practices. This Act also establishes research institutions such as the Kenya Maritime and Fisheries Research Institute and Kenya Medical Research Institute with the mandate to carry out research on marine life and human health respectively.

The National Guidelines for Ethical Conduct of Biomedical Research Involving Human Participants in Kenya These guidelines, issued by NACOSTI, provide ethical standards for biomedical research in Kenya.

Copyright Act which allows for the restriction of copyright in research through fair dealing. Fair dealing is a legal exemption that enables users to reproduce or utilize copyrighted material without permission or payment to conduct scientific research.

The Industrial Properties Act restricts the submission of patent applications for research that would jeopardize the safety of the public and/or the defense of the country if published. The Act also exempts scientific research from patent rights protection and prohibits terms in license contracts that contain provisions that seek to restrict research or technological development to improve existing products.

The Universities Act establishes the Commission of University Education that has the mandate to promote quality research and innovation. The Act also provides that universities should promote quality and ethical standards of carrying out academic research.

The Science, Technology and Innovation (Research Licensing) Regulations, 2014 This regulation obligates all persons intending to undertake scientific research in Kenya to obtain a license from NACOSTI in accordance with the STI-Act.

Pharmacy and Poisons Act provides regulation on research on poisons, drugs and controlled substances. The Act also provides a basis for clinical trials through the **Pharmacy and Poisons (Conduct of Clinical Trials) Rules 2022.** The Rules provide detailed guidelines on seeking approvals and requirements for persons/institutions carrying out clinical trials.

6.1 Lawfulness, Fairness & Transparency

Researchers must process personal data lawfully, fairly and in a transparent manner. Lawfulness means that any collection, use, or processing of personal data must be based on valid legal grounds. Fairness means that researchers should handle personal data proportionally and in ways that do not cause harm to data subjects. In processing personal data fairly, a researcher should consider how it affects the interests of the people concerned either as a group or an individual. Transparency in processing personal data requires that the individuals concerned are informed about who the researcher is, how and why they will use their personal data. A researcher should inform concerned people about their processing in a way that is easily understandable.

Example:
Msomi University was conducting research on alcohol and substance abuse in institutions of higher learning. The university established a clear legal basis for processing personal data, as outlined in the Act which includes obtaining explicit consent from participants, as research on sensitive topics like alcohol and substance abuse requires a strong justification for data collection. Participants were provided with comprehensive information about the research, including the objectives, procedures, and potential risks. This allows participants to make an informed decision about their involvement without any coercion. The university clearly communicated the specific data being collected, such as demographic information, substance use patterns, and attitudes toward alcohol and drugs. This transparency helps participants understand the scope of the study and the types of data they are providing. Lastly the University explained the safeguards in place to protect personal data which include secure storage and data anonymization.

6.2 Purpose Limitation

Purpose limitation is a crucial aspect of data protection. It requires researchers to document and regularly review the alignment of collected data with the stated research purposes. Personal data collected for research must not be repurposed for commercial or unrelated government functions without renewed legal basis and appropriate safeguards.

Example:
While undertaking a master's research project, Mwerevu conducted research on the profitability of small-scale farmers in rural region of Kenya. The purpose of the research was purely for academic purpose and to enable her complete her masters studies. The personal data processed for purposes of this research was therefore, limited purely to academic research.
Upon completion of her degree, Mwerevu started a Company selling tools to assist small scale farmers improve their productivitity. She cannot use the personal data collected for her academic reserach to directly market the tools.

6.3 Data Minimization

Data minimization is a critical aspect of data protection when processing personal data for research purposes. The principle requires that personal data collected and processed is relevant and limited to what is necessary for the research purpose. Researchers should avoid collecting excessive data that has nothing to do with the purpose of research and should not collect broad data in anticipation of unforeseen purpose.

Example:
The State Corporation in charge of Transport is researching commuting patterns to improve public transportation services in a city. The team decides not to collect names or any other personal identifiers of commuters. Since the purpose is to analyze patterns rather than individual behaviors, collecting names is unnecessary and poses privacy risks. Instead of gathering specific arrival and departure times, the researchers use generalized timing data (morning, midday, evening). This allows them to identify peak commuting times without needing exact details that could link back to individuals. The survey includes only essential demographic questions, such as Age range (18-25, 26-35), mode of transport used (bus, train, walking). This limited demographic information supports the analysis while minimizing personal data collection. Through its data minimization strategy, the State Corporation effectively gathers valuable insights into commuting patterns while safeguarding the privacy of commuters.

6.4 Accuracy

Accuracy is to ensure that personal data for research purposes is correct, reliable and up to date, supporting the validity of findings and research outcome. A researcher must take

reasonable steps to verify the correctness of personal data and ensure that inaccurate personal data is rectified or deleted. This includes verifying data at the point of collection and periodically checking it for relevance and correctness throughout the research process to ensure high quality of data.

Further, you should incorporate robust data accuracy practices into their methodologies, balancing legal requirements with practical research needs.

A research institute is conducting a survey on socio-economic status in different counties to inform policy decisions. They collect data on income level, employment status, and educational background. The institute must use a reliable source for data collection, implement processes to verify data accuracy, periodically review and update data and establish where possible for participants to report inaccuracies and request corrections.

6.5 Confidentiality & Integrity

Confidentiality and integrity mean that personal data used in research must be accurate, complete, and protected from unauthorised alteration. In research, applying confidentiality and integrity principles is crucial to ensure ethical standards, enhance the credibility and reliability of the research findings and protect participants privacy. In research, applying confidentiality and integrity principles is crucial to ensure ethical standards, enhance the credibility and reliability of the research findings and protect participants privacy. Security measures to uphold the confidentiality and integrity of data must be in place before the start of the research and throughout the research process. For example, any hard copies of notebooks, questionnaires, surveys and other paper documentation should be kept in a secure location, where there is no public access.

Example
Mwalimu, an educator is conducting a study on the effects of social media on mental health among young adults. She has designed a survey to gather data from participants and has obtained informed consent from each participant. To ensure the confidentiality and integrity of the data she takes the following measures:
a) Anonymization of Data; No Personal Identifiers: Mwalimu collects responses without any identifiable information, such as names or contact details. Instead, participants are assigned unique identifiers like id numbers to protect their identities. The findings are reported in aggregate form, ensuring that individual responses cannot be traced back to any participant.
b) Secure Data Storage; Encrypted Data Collection: While using online surveys, Mwalimu ensures that the platform is secure and data is encrypted during transmission and storage.
Physical Security: For paper surveys, Mwalimu stores completed questionnaires in a locked filing cabinet, accessible only to authorized personnel.
c) Access Controls: Mwalimu restricts access to the data to only those directly involved in the research. Access logs are maintained to track who has accessed the data.
d) Research Ethics Training: Mwalimu ensures that all team members involved in data handling are trained on ethical research practices and data protection protocols.
e) Retention Policy: Mwalimu establishes a clear policy for how long data will be retained and ensures that data is only kept for the necessary duration.
f) Secure Disposal: After the retention period, any physical copies of data are shredded, and digital files are permanently deleted from all devices and storage systems or data anonymized.
g) Training of team members: Mwalimu has conducted a training for the team members on how to handle data during throughout the research process from collection, analysis of data and reporting as well as security and confidentiality of data.

6.6 Storage Limitation

This principle ensures that personal data is not kept for longer than necessary for research purposes, while also addressing ethical considerations and compliance with legal requirements. Section 25 (g) of the Act requires that every data controller or data processor shall ensure that personal data is kept in a form which identifies the data subjects for no longer than is necessary for the purposes. Storage limitation means that, even if you collect and use personal data fairly and lawfully but you no longer need it, you cannot keep it for longer than you need it. Researchers must maintain a data retention schedule, including justification for extended retention and periodic review mechanisms.

Example
A health research team at a university conducted a study on malaria prevention practices among rural homes in Western Kenya. The study was completed three years ago, and all analysis, reporting, and publication were finalized. However, the team continues to store personal data, including names, home geolocations, and health records, "just in case" it may be useful for future projects.
Upon internal review triggered by the institution's updated data protection policy, it was noted that:
i) The study's original ethics approval and consent forms did not cover future use or long-term retention;
ii) The data was not stored in encrypted folders on shared drives;
iii) No justification had been documented for the extended storage.
Recognizing the violation of the storage limitation principle, the team took the following steps:
i) Conducted a retention review and found that the data was no longer necessary for the completed research.
ii) Anonymized aggregated results already published were retained for archival purposes.
iii) Securely deleted identifiable records from all storage locations.
iv) Updated internal policies to ensure future studies include clear data retention schedules and anonymization protocols.

However, section 39 of the Act provides limitation to the principle of storage limitation for research-related processing. It permits retention of personal data if the data is required for historical, statistical, journalistic, literature and art, or research purposes to ensure comprehensive and meaningful results.

Nevertheless, if personal data are no longer required for the purpose of research, a data controller or processor must delete, erase, anonymize, or pseudonymize personal data that is no longer needed.

Guidance:
In research, long-term data retention is often necessary to ensure comprehensive and meaningful results. Section 39(1)(d) allows for this extended retention while ensuring compliance with the overall intent of data protection.

6.7 Accountability

The principles of accountability emphasize responsible data handling practices, compliance with legal requirements, and respect for the rights of data subjects. This involves taking

proactive steps to protect personal data. The accountability principle in processing of personal data for research purposes refers to the responsibility of researchers to ensure compliance with relevant laws and regulations related to the collection, use, and disclosure of personal data for such processing.

As a researcher you should have policies and procedures in place to govern the collection, use, and disclosure of personal data. These policies and procedures should be reviewed and updated regularly to ensure compliance with changing legal and regulatory requirements.

Example:

Majibu research institute is conducting a study on the effects of climate change on agricultural practices among smallholder farmers. The study involves collecting personal data from farmers, including demographic information and farming practices.

The institution establishes a clear data protection policy that outlines responsibilities regarding the collection, processing, and storage of personal data. This policy complies with the- Act and is publicly available.

All staff involved in the research receive training on data protection, emphasizing the importance of accountability. This includes understanding their responsibilities for safeguarding personal data and the ethical implications of data handling.

Prior to data collection, the researchers obtain informed consent from all participants. Consent forms detail the purpose of the research, how data will be used, and the participants' rights regarding their information.

The research team maintains comprehensive records of data processing activities, including, types of data collected, the legal basis for data processing, data retention periods and who has access to the data.

After the study concludes, the research team evaluates its data protection practices including whether data retention practices were followed and securely delete any personal data that is no longer needed.

This approach not only ensures compliance with the Act but also enhances trust among participants, fostering a responsible research environment.

LAWFUL BASIS FOR PROCESSING DATA FOR RESEARCH PURPOSES

The Act provides for several lawful bases for processing personal data. Section 30 of the Act requires that personal data shall only be processed if at least one of the several legal grounds listed in that Section applies.

One of the bases for processing is scientific research. While processing for purposes for scientific research, there are conditions to be met in the Science Technology Innovations Act. In practice, the original basis to collect the data may not always be appropriate for your research-related processing which means research cannot be used as a standalone lawful basis. The lawful basis for research-related processing can either be:

- Consent
- Performance of a task carried out by a public authority
- Legitimate interests
- Public interest
- Scientific research

7.1 Consent

Consent is the cornerstone of any research with human participants. Section 32 of the Act provides that Consent must be freely given, informed, specific and unambiguous. (Refer to Guidance Note on Consent https://www.odpc.go.ke/guidelines-2/). Consent gives participants real control over how their data is used. Data subjects have the right to withdraw their consent at any time and have their data deleted. Deletion of data is limited to the personal data itself and not the product of research.

Where consent is relied upon, it must be informed, specific, and freely given, and capable of being withdrawn without detriment to the data subject. Consent mechanisms developed in relevant sectors can be relied upon but should meet the conditions of consent set out under Section 32 of the Act.

Before administering consent, a researcher is required inform the research participant of the following but not limited to:

i. Invitation to participate in the research;
ii. Title of the project;
iii. Principal investigator(s) name(s);
iv. Institution undertaking research;

v. Type of personal data to be collected;
vi. Purpose of the research study;
vii. Role of study participants and duration of participation;
viii. Research methodology;
ix. Benefits of the research project;
x. Unforeseen risks/discomfort;
xi. Confidentiality assurances especially on identifiable records;
xii. Investigators responsibility;
xiii. Storage of personal data both general data and sensitive data;
xiv. Non-coercive disclaimer /voluntariness to participate and to withdraw at any time and the consequences of withdrawing consent; and
xviii. Dissemination of findings.

The researcher should ensure that they provide detailed information in a transparent manner to the research participant to enable them to make an informed decision on whether to give consent or not.

Example
Mwerevu, a PHD student, undertaking a study on the effects of the COVID-19 pandemic on women in Kenya has acquired a license to undertake this study from NACOSTI. Her study is informed by real life experiences of women in Kenya. In order to collect these experiences, she has issued consent forms to each of the participants. The consent form contains; the purpose of the study, the process of the study, risks associated with giving personal information, statement of confidentiality, information on how to withdraw such consent, measures to safeguard personal data from aunathorized access, how the findings will be disseminated and the publication of findings.

7.2 Performance of a Task Carried out by a Public Authority

This basis would apply for Public Authorities such as Public Universities, Government agencies and research council institutes to conduct research pursuant to an Act of parliament, Legal Notice or Executive Directive.

7.3 Legitimate Interest

Example
Ujuaji Medical Research Institute is established by a Legal Notice. It has a legal mandate to monitor, assess, and improve the health of its residents. As part of this mandate, the department aims to conduct a research study to understand the relationship between air pollution levels and various health outcomes in urban areas. This research will involve processing personal health data and environmental data.

Legitimate Interest is the most appropriate legal basis for research done by organizations, such as commercial enterprises and nonprofit research institutes that are not part of the government. Legitimate interest allows organizations to process personal data if it is necessary for their legitimate interests or those of a third party, provided that these interests are not overridden by the rights and interests of the individuals whose data is being processed.

Organizations must clearly identify and articulate their legitimate interests in conducting research. This could include advancing scientific knowledge, improving products or services, or addressing social issues.

A researcher must examine the necessity and proportionality of processing sensitive personal data. Researchers relying on legitimate interest are required to conduct a balancing test to weigh their legitimate interests against the rights and freedoms of individuals. If the risks to individuals' privacy and data protection rights outweigh the benefits of the research, the legitimate interest basis may not be appropriate.

7.3.1 Legitimate Interest Assessment Framework

Determining whether the processing of personal data for research purposes for legitimate interest involves a thorough assessment. This framework can guide data controllers in making that determination:

a) **Define the Purpose and Benefits**

Clearly define the specific purpose of the data processing and demonstrate the legitimate interest being pursued. This should include a rationale for why the processing is necessary and how it benefits the organization, the public, or a specific group while ensuring it does not override the rights and freedoms of the data subjects.

Example:
A financial firm analyses transaction patterns to detect emerging fraud risks among small business clients. The legitimate interest is to enhance platform security and protect users without infringing on customer privacy.

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:
A public hospital wants to evaluate the effectiveness of a maternal health outreach program in rural counties. The analysis requires patient visit records, age, pregnancy history, and treatment outcomes over a 3-year period. To ensure necessity and proportionality:
i) The hospital removes names and national ID numbers (pseudonymization);
ii) Uses the data solely for internal performance evaluation;
iii) Limits access to a small, authorized research team; and
iv) determines that surveying new patients alone would not yield accurate retrospective insights.

c) **Implement Safeguards and Mitigations**

Identify and implement appropriate safeguards to protect data subjects' rights and privacy. This includes data anonymization, access controls, data minimization, and secure data handling practices.

Example:
The data controller anonymizes the health records and limits access to a small group of researchers. Data is stored securely, and regular audits are conducted to ensure compliance with data protection standards.

d) **Consider Public Expectations and Transparency**

Consider 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 organization 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 Legitimate Interest Against Potential Harm

Conduct a balancing test to weigh the legitimate 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:
A local university conducts a survey among its students to understand patterns of library use, including preferred study hours, most accessed resources, and frequency of visits. The university relies on legitimate interest to process this data being the interest of improving library services, optimizing resource allocation, and enhancing student experience.

f) Document the Assessment

Maintain comprehensive records of the legitimate 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.

7.4 Public Interest

Research for the purposes of public interest aims to generate knowledge beneficial to society, addressing issues affecting welfare, policy, and communal well-being, often focusing on public health, education, environment, and social justice. When undertaking such research, organizations can rely on the public interest as a lawful basis.

Example:
Health for All, a non-governmental organization (NGO), is conducting a study on the impact of water quality on public health in rural communities. The research aims to inform policymakers and improve access to clean water, which aligns with public health objectives and community welfare. Health for All, has been collaborating closing with the Ministry of Health of Kenya in furtherence of public heath policies. The public interest in this research is to enhance public health outcomes by identifying how contaminated water sources contribute to diseases in rural areas. The findings will be used to advocate for better water quality management and resource allocation.
The NGO collects personal data from participants, including demographic information, health records, and details about water sources they use. This data is essential for understanding the relationship between water quality and health outcomes.
The processing of personal data in this example is necessary to achieve the research objectives. Without this data, the organization would be unable to draw meaningful conclusions that could lead to actionable public health interventions.
By focusing on the broader benefits to community health and welfare, the NGO used Public Interest to collect and use personal data while ensuring compliance with the Act.

7.5 Scientific Research Purposes

When relying on the lawful basis of Scientific research, researchers need to adhere to the relevant conditions set out in the Science, Technology and Innovation Act. According to Part IV, there are several requirements to meet to be licensed to undertake research. Generally, a person may undertake scientific research in Kenya after obtaining a licence under the Science, Technology and Innovation Act, unless the scientific research is to be conducted within a university or an academic research institution programme. Also, any person undertaking or intending to undertake research in science and technology in the country, or who accesses, handles, or transfers any material or technology or moves it within, from or into the country, shall apply to the National Council for Science and technology for a licence in accordance with the Science, Technology, and Innovation Act.

PROCESSING OF SENSITIVE PERSONAL DATA

Section 44 of the Act states that sensitive data must be processed in compliance with the principles of data protection. Processing of sensitive data requires additional safeguards. A researcher must examine the proportionality and necessity of processing sensitive personal data and must also identify one of the specific conditions for processing sensitive personal data as outlined in Section 45 of the Act. A Data Protection Impact Assessment (DPIA) must be conducted where research involves large-scale processing of data or sensitive data related to vulnerable populations.

Best Practice
Seek informed consent before any processing of sensitive Personal Data. Section 2 of the Act provides that Consent must be freely given, informed, specific and unambiguous.

8.1 Clinical Trials

Clinical Trials involve human participants and are regulated by the Pharmacy and Poison Board through the Guidelines for the Conduct of Clinical Trials in Kenya and Pharmacy and Poisons (Conduct of Clinical Tests) Rules. Research data in clinical trials contains sensitive data such as information relating to clinical research participants medical condition, medications they are taking or have taken in the past, laboratory values and reports, medical history, and more. As such, this data is subject to the principles of data protection.

It is mandatory to obtain informed consent for clinical trials and should ensure that all stakeholders engaged such as sponsors, (Clinical Research Organisations) CROs and hospitals/doctors must adhere to all privacy procedures and requirements. That ensures that the bodily autonomy of a trial's human test subject is protected. Research involving biometric, genetic, or health data must undergo independent ethical review and incorporate additional safeguards, including data minimization and pseudonymisation.

8.2 Assessing Vulnerability of Research Participants

Vulnerable research participant means a person who needs special care, support or protection, including the orphaned and vulnerable children, widows or widowers, person with disability, elderly persons economically or educationally disadvantaged individuals, people in institutional settings (e.g., prisoners or hospital patients), cultural or ethnic minorities among other groups or indigent due to a risk of abuse or neglect. They are those persons who are relatively (or

absolutely) incapable of protecting their own interests. More formally, they may have insufficient power, physical capacity, intelligence, education, resources, strength, or other needed attributes to protect their own interests.

Processing personal data when dealing with vulnerable research participants requires careful consideration and strict compliance with the data protection principles. The lawful basis used to process personal data from vulnerable research participants must be tailored to address the specific vulnerabilities of the participant. The researcher must demonstrate that an assessment on vulnerability was done before the research began and demonstrate how the vulnerability risks are to be addressed during the personal data processing.

The Office guides that vulnerability assessment of research participants especially vulnerable populations in the society is necessary since these groups are open to physical, social, economic, physical, emotional and psychological exploitation. Vulnerability assessment should be done in line with sector specific guidelines on vulnerable populations while ensuring the right to privacy of vulnerable research participants.

In conducting vulnerability assessment, a researcher is required to:

i. Contextualize the vulnerability whether its cognitive/communicative, social, economic, physical, emotional etc. and how the identified risks will be addressed;
ii. Assess the ability to make an informed decision including voluntariness by the vulnerable participants;
iii. Identify appropriate lawful basis e.g. for consent, use tailored consent processes without undue coercion or inducement;
iv. Assess the necessity vs proportionality while justifying the inclusion of the vulnerable participants and how they will be protected so that the vulnerabilities will not be exacerbated;
v. Outline mechanisms to ensure adherence to all the data protection principles of fairness, transparency, purpose limitation, data minimization, accountability, confidentiality and integrity, accuracy and storage limitation;
vi. Outline mechanisms on how vulnerable participants can exercise the right to withdraw consent, access, correction, and deletion; and
vii. Provide organizational and technical safeguards to safeguard personal data from unauthorized access and misuse. A researcher should identify any additional

safeguards that may be necessary when handling personal data from vulnerable research participants.

9. COMPLIANCE OBLIGATIONS

All entities or individuals participating in research should comply with the following obligations:

a) Registration with the Office as data controllers and or data processors;
b) Assess eligibility to conduct research defined by applicable laws or regulations under Science, Technology and Innovation Act under NACOSTI;
c) Taking a 'data protection by design and default' approach to your processing activities;
d) Implementing appropriate technical and organization security measures;
e) Establish data retention and disposition schedule with clear procedures for destruction. If the researcher researchers/ organisations no longer processing the data for any purpose, including a research-related purpose, a researcher must delete it, can also anonymize and pseudonymize personal data;
f) Carrying out DPIAs on all their processing activities that are likely to result in a high risk to the rights and freedoms of data subjects and submit the data impact assessment reports to Office where necessary and should be update them throughout project lifecycle. Refer to Guidance Note on DPIA https://www.odpc.go.ke/guidelines-2/);
g) Appointing a data protection officer DPO), where necessary;
   A DPO is particularly required when:
   i) The processing involves large-scale sensitive personal data, such as health or biometric data in clinical trials;
   ii) The research involves systematic monitoring of data subjects, such as behavioural or observational studies using tracking technologies;
   iii) The organization is a public authority or body; or
   iv) The processing is high-risk, especially where it involves vulnerable populations or could significantly impact data subjects' rights and freedoms.
h) Providing regular and comprehensive data protection appropriate trainings for all the staff handling or processing personal data;
i) Notification of data breaches to the Data Commissioner within 72 hours of occurrence and communicating with the data subject in writing within a reasonably practical period, unless the identity of the data subject cannot be established;
j) Assessing and documenting the lawful basis for regular data sharing with third parties and entering into data sharing agreements;

k) Use of privacy-enhancing technologies e.g. trusted research environments; and

l) Obtaining consent from data subjects when there is transfer of sensitive data.

10. RIGHTS OF DATA SUBJECTS

Right	What does this mean?	Are there limits to this right?
Right to be informed	Right to be provided with extensive information including on 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 calendar 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 would adversely affect 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 researcher 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, • 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.	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 related public health, • fulfilment of legal claims, • 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 of 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 day(s).	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.	Subjects have the right to withdraw consent at any time and researchers must inform them of this right prior to data collection. Withdrawal of consent shall not affect the lawfulness of processing based on prior consent before its withdrawal. 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 decision making	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.

10. EXEMPTIONS

10.1 Conditions for Exemptions

There are no absolute exemptions under the Act and a researcher should be able to justify the exemptions on a case-by-case basis.

Further, the researcher should ensure that:

a) the research carried out involves activities that are beneficial to the society at large, communities in Kenya, environment;
b) the research should not adversely affect the lives of Kenyans.

Personal data which is processed only for research purposes is exempt from the provisions of the Act if it meets the following criteria:

a) The data is processed in compliance with the relevant conditions set out in Science Technology and Innovation Act, and;
b) The results of the research that are made available to others cannot be used to identify the data subjects the information was collected from.

The above means that the researcher cannot be self-proclaimed, but there is a strict procedure to implement. Scientific research should meet the conditions set out in the Science Technology and Innovation Act. Another important facet of this exemption is that the results of the research are presented in a manner that makes the identification of the data subject impossible. In other words, the researchers are required to keep the identity of the persons involved confidential. Any exemption must be approved by the ODPC and based on a formal justification that the research cannot be conducted using anonymized data and serves a substantial public interest. The research output should, however, not be published in an identifiable form.

10.2 Exemptions under Public Interest

Research carried out for public interest purposes is exempt if it falls under permitted health situations. However, researchers carrying out research under a permitted health situation should provide a clear lawful processing, ensure minimisation of collection, data quality, and demonstrate mechanisms of adopted to implement organizational and technical security safeguards to protect personal data as required under Section 51 (1) of the Act. A permitted

health situation involves the collection, use, or disclosure of health data for health research and related purposes for public interest, for example during a pandemic.

Guidance:
In the event of a pandemic, public health authorities might research the prevalence of the disease across various age groups. This research would be exempt from the Act, as it qualifies under permitted health situations exempted in the public interest. However, the relevant authority would still need to meet the requirements set by the Science, Technology, and Innovation Act.

10.2.1 Public Interest Assessment Framework

Determining whether the processing of personal data for research 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 public or the larger society. Research benefits should not be limited to a specific group or individuals. This includes understanding how the research or analysis will contribute to public knowledge, policy-making, or cultural preservation

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.

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 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.

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.

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.

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.

10.3 Exemptions on further processing for Compatible Purposes

Further processing of personal data for research purposes is considered a compatible purpose and is permitted on condition that is carried out solely for such purposes and it is not published in an identifiable form. In this case, a data controller or data processor does not need to seek fresh consent from data subjects. However, a data controller or data processor should put in place appropriate technical and organisational measures to safeguard against the records being used for any other purposes and the results should not be published in an identifiable form.

To determine compatibility, data controllers and processors should assess whether the further processing:

i) Aligns with the original research-related purposes;
ii) Is necessary, proportionate, and fair;
iii) Is carried out lawfully; and
iv) Is supported by appropriate technical and organisational safeguards, such as anonymisation or pseudonymisation.

A data controller or data processor should ensure that safeguards are in place to prevent the data from being used for unrelated purposes or disclosed in a manner that identifies the data subjects.

Example 1:
Msomi University collects personal data to register new students and continually provide them with necessary resources. Later the university undertake a research program to identify weaknesses in their system and discover more effective ways to support their students.
When Compatible: The university can utilize the personal data held for this research as it is compatible with the original purpose of enhancing student resources and improving overall services. The research aims to benefit the same group of individuals (students) for whom the data was initially collected.
When Not Compatible: If Msomi University were to use the personal data for unrelated purposes such as marketing products unrelated to education or sharing data with third-party commercial entities. It would not be considered compatible with the original purpose of registration and resource provision. In such cases, explicit consent from the data subjects would be required to ensure compliance with the Act and Regulations.

Example 2:
XYZ financial institution relies on legitimate interest and collects personal data from sampled individuals that they wish to use to test a product or service before launching it. The personal details collected could include names, phone numbers, location, income levels/property details etc.
After this product has been launched, the financial institution may carry out research using the above individuals to understand how the product or service is benefiting them without being required to obtain a new consent as the purpose of the research is compatible with the initial purpose.

11. SECURE DATA HANDLING WHILE CONDUCTING RESEARCH

Researchers must implement appropriate technical and organizational measures to protect personal data from unauthorized access, loss, alteration, or destruction. Technical measures could include anonymisation, pseudonymisation, encryption, access controls, regular security audits, and implementation of data protection by design and default. Organizational measures could include organizational policies on data security, training of employees processing personal data, secure handling of physical files, use of CCTVs and lockable doors and cabinets for storage, data storage and retention schedules with appropriate disposition mechanisms etc. These safeguards minimise the risk of re-identification, unauthorized access or breaches and protect privacy of individuals throughout the research process.

11.1 Anonymisation and Pseudonymisation

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.

11.1.1. Anonymisation

Anonymisation is the process of removing or modifying personal data so that individuals cannot be identified, directly or indirectly, by any means reasonably likely to be used. Before data is anonymized, researchers should ensure strict adherence to the data protection laws. Once data is anonymized, it falls outside the scope of data protection laws, as it no longer constitutes personal data. Researchers are however required to use techniques that ensure that anonymisation is irreversible.

Steps for Effective Anonymisation:

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 anonymisation**: Assess the data to ensure that individuals cannot be re-identified using reasonably available information.

Example:
A university conducting research on population health anonymises patient records by removing names, addresses, and exact birthdates, and generalizes age into ranges (e.g., 20 - 30, 31 - 40). The removal of the personal identifiers is done by permanent deletion. This ensures that the data can be used for research without risking the identification of individuals.

11.1.2. Pseudonymisation

Pseudonymization is the process of replacing private identifiers with fake identifiers or pseudonyms. Unlike anonymisation, 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 Pseudonymisation:

i. **Replace Identifiers**: Substitute identifiers with pseudonyms (e.g., replace names with unique codes).

ii. **Separate Data**: Store the pseudonymisation 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.

11.2. Non-Identifiable Data for Publication of Research Output

Section 53(3) prohibits publishing personal data processed for research purposes in an identifiable form. Results must be anonymized or aggregated 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 University uses aggregated data of students using the University library to show trends without revealing individual identities. If the study needs to include specific personal experiences and satisfaction on the library services from the students, the University must obtain explicit consent from those individuals.

APPENDIX A: COMPLIANCE CHECKLIST FOR INTERNAL USE BY RESEARCHERS.

No.	Compliance Requirement	Questions	Yes	No	Comments/Remedial Actions
1	Lawful basis for processing	Are you eligible to conduct the research as per the specific laws and regulations?			
		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 the processing purposes and the lawful basis in our privacy notice.			
		Where we process sensitive data, we have also complied with section 44, 45 and 46 of the Act			
2	Compliance Obligations	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 conducted a DPIA?			

		Have you conducted a vulnerability assessment of research participants?			
		Have you adopted technical and organizational security measures?			
		Are you sharing sensitive personal data with third parties? If so, have you entered into formal data sharing agreements with the third parties to specify the responsibilities and safeguards for data protection?			
3	Exemptions	We have complied with other non-exempt principles listed in section 51 of the Act. (Lawful processing, data minimization data quality, and safeguards)			
		Our research falls under relevant conditions set out in the Science, Technology & Innovation Act			
		We have put in place measures for safeguarding personal data in the research			
		Further processing of personal data is purely for research, historical and scientific purposes			
		The research results will be published in a manner that does not identify individuals			
		Our research does not have adverse effects on the participants and the society at large			
4					