National Research and Development Institute

RIKEN

Guidelines for Personal Information Management in Medical Research Involving Human Subjects

March 26, 2020

Committee for ICT Infrastructure Development
and Service Cooperation

TABLE OF CONTENTS

1. Objective, etc. of the Guidelines

1-1. Objective of the Guidelines

The Guidelines specify matters to be observed by the researchers of RIKEN, national research and development institute, (hereinafter the "Institute") concerning management of personal information in medical research involving human subjects based on the requirements of the following two documents issued by the Ministry of Education, Culture, Sports, Science and Technology and the Ministry of Health, Labour and Welfare: the *Ethical Guidelines for Medical and Health Research Involving Human Subjects* (established on December 22, 2014 and partly amended on February 28, 2017; hereinafter the "Ethical Guidelines"[1]) and the *Guidance on Ethical Guidelines for Medical and Health Research Involving Human Subjects (full edition)* (established on February 9, 2015 and partly amended on May 29, 2017; hereinafter the "Guidance" [2]), which provides interpretations of the Ethical Guidelines and specific points to note in the procedure stipulated in the Ethical Guidelines.

The Institute shall prevent leakage, loss, or damage of personal information that it holds. Specifically, medical data and genome data used for medical and health research involving human subjects, contains highly delicate information requiring special care. If such information should be leaked, lost, or damaged even on a small scale, the resulting impact would be enormous.

Therefore, researchers of the Institute shall handle personal information and other private information properly in accordance with relevant laws and regulations as well as the Guidelines.

1-2. Laws, Regulations, and Policies Pertaining to the Guidelines

Personal information held by the Institution shall generally be handled in accordance with the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc.[3] (Act No. 59 of 2003; hereinafter the "Personal Information Protection Act").

According to the Personal Information Protection Act, incorporated administrative agencies shall not, in principle, use or provide another person with personal information or anonymously processed information held by them for purposes other than the purpose agreed to by the data subject.[4][5]

There are, however, cases where such information may be used or provided to another person for purposes other than the agreed purpose of use. Unless otherwise specified in other laws or regulations and except for cases where the purpose of use is likely to cause unjust harm to the rights or interests of the data subject or a third party of the personal information, the Institute may use retained personal information internally only to the extent necessary for executing the business provided by laws and regulations and when there are reasonable grounds for the use of that retained

---

[1] https://www.mhlw.go.jp/file/06-Seisakujouhou-10600000-Daijinkanboukouseikagakuka/0000153339.pdf
[2] https://www.mhlw.go.jp/file/06-Seisakujouhou-10600000-Daijinkanboukouseikagakuka/0000166072.pdf
[3] https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000129084.pdf
[4] Article 4 of the Personal Information Protection Act
[5] Articles 9.1 and 44.2.2 of the Personal Information Protection Act

personal information or provide such information to another person exclusively for academic research purposes.[6] The term “academic research” here refers to finding new laws or principles, establishing analytical and methodological theories, systematizing new knowledge and its applications, and developing advanced fields of study. It shall also be noted that, among the research conducted at the Institute, research where personal information is provided for the purpose of developing products with an external joint research institution is not regarded as “exclusively for academic research purposes.”

When personal information (including information on the deceased and the unborn; the same shall apply hereinafter) is used for medical and health research involving human subjects (hereinafter the “Research”), researchers shall use personal information properly and securely, respecting the human dignity and rights of the research subject in accordance with the above-mentioned basic framework of the Personal Information Protection Act and under the academic freedom guaranteed by the Constitution. Specific measures are stipulated in the Ethical Guidelines (the *Ethical Guidelines for Medical and Health Research Involving Human Subjects*) and the Guidance (the *Guidance on Ethical Guidelines for Medical and Health Research Involving Human Subjects (full edition)*.[7]

The Guidelines, established in accordance with the requirements of the Ethical Guidelines and the Guidance, specify matters concerning secure management of personal information in research activities that shall be observed by the researchers of the Institute. Even if other guidelines[8] on research set forth requirements regarding secure management of personal information, the researchers are required to abide by the Guidelines wherever possible.

Requirements for secure management of personal information in the Guidelines were formulated based on the *Security Guidelines for Medical Information Systems, 5th Edition* (issued in May 2017

---

[6] Articles 9.2.2, 9.2.4, and 9.3 of the Personal Information Protection Act

[7] Though the Ethical Guidelines and the Guidance are not legislation, failure to observe them is deemed as misconduct, possibly resulting in a return of research funds, as described in the *Guidelines on Research Misconduct* (approved on August 26, 2014 by the Minister of Education, Culture, Sports, Science and Technology).

[8] The *Ethical Guidelines for Human Genome / Genetic Analysis Research* (instituted on March 29, 2001 and partially amended on December 1, 2008 by the Ministry of Health, Labour and Welfare); the *Guidelines for Clinical Research Such as Gene Therapy* (instituted on October 1, 2015 and partially amended on February 28, 2019 by the Ministry of Health, Labour and Welfare); the *Discussion about the Approach of Research and Development Using Human Tissues Obtained from Surgery* (issued on December 16, 1998 by the Ministry of Health, Labour and Welfare); the *Basic Guidelines for the Conduct of Animal Experiments in Implementing Agencies Under the Jurisdiction of the Ministry of Health, Labour and Welfare* (instituted on June 1, 2006 and partially amended on February 20, 2015 by the Ministry of Health, Labour and Welfare); the *Guidelines on Issues of Infectious Diseases in Public Health Associated with Clinical Trials of Xenotransplantation* (issued on July 9, 2002 and amended on June 13, 2016 by the Ministry of Health, Labour and Welfare); the *Ethical Guidelines for Research on Assisted Reproductive Technology Treatment Producing Human Fertilized Embryos* (instituted on December 17, 2010 and partially amended on February 28, 2017 by the Ministry of Health, Labour and Welfare and the Ministry of Education, Culture, Sports, Science and Technology); the *Ethical Guidelines for Epidemiological Research*, (instituted on June 17, 2002 and partially amended on December 1, 2008 by Ministry of Health, Labour and Welfare); the *Ethical Guidelines for Clinical Research* (instituted on July 30, 2003 and wholly amended on July 31, 2008); the *Guidelines for Clinical Research Using Human Stem Cells* (instituted on July 3, 2006 and abolished on November 24, 2014 by the Ministry of Health, Labour and Welfare)
（https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hokabunya/kenkyujigyou/i-kenkyu/index.html）

by the Ministry of Health, Labour, and Welfare, the *Safety Management Guidelines for Information Processing Businesses Handling Healthcare Information on Behalf of Others, 2nd Edition* (issued in October 2012 by the Ministry of Economy, Trade and Industry), and the *Guidelines Relating to Safety Management When Cloud Service Businesses are Handling Healthcare Information, 1st Edition* (issued in July 2018 by the Ministry of Internal Affairs and Communications).

Other relevant laws and regulations include the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc.*[9] (issued in March 2017 by the Personal Information Protection Commission), and the *Measures for the Proper Management of Personal Information Held by Administrative Organs* (issued on September 14, 2004 and finalized on October 22, 2018 by Director-General of the Administrative Management Bureau of the Ministry of Internal Affairs and Communications).[10] Provisions on the protection of personal information in clinical trials are included in the Act on Securing Quality, Efficacy and Safety of Products Including Pharmaceuticals and Medical Devices (Act No. 145 of 1960), and those on the provision of medical data to external parties in the Act on Anonymized Medical Data That Are Meant to Contribute to Research and Development in the Medical Field[11] (Act No. 28 of 2017; hereinafter the "Next Generation Act").

Researchers are required to conduct research properly by referring in a timely manner to the above-mentioned laws and other rules relevant to research as well as the Guidelines.

1-3. Premises of the Guidelines

The Guidelines specify necessary and proper measures[12] that researchers shall take to protect personal information held by the Institute. The measures include means of protecting personal information from being leaked, lost, or damaged.

First and foremost, the Institute and its researchers shall fully understand that the personal information to be used for research or obtained through research does not belong to researchers, but to the Institute

---

[9] https://www.ppc.go.jp/files/pdf/guidelines06.pdf

[10] http://www.soumu.go.jp/main_content/000579982.pdf

[11] Under the Personal Information Protection Act, clinical data of an individual held by a private medical institution may be provided to a third party, such as a corporate or research institution, either upon consent of the individual (opt-in system) or after the data has been anonymously processed by the institution at its own responsibility or an external information processor entrusted by the institution in such a way that the data can no longer be attributed to the individual. Through the enactment and enforcement of the Next Generation Act, medical institutions are now able to provide raw data to a state-accredited, anonymously processed information handling business operator based on the opt-out method, which allows the institutions to do so, unless expressly rejected by the individual at their first visit. This is expected to help medical institutions gather high-quality medical data and build a large database using them.

[12] These measures are aimed to develop a system for secure data management as stipulated in Article 6.15 of the Ethical Guidelines.

(including in the case of entrusted information). Personal information is to be held by the Institute and managed properly under its supervision.

The Guidelines also specify the management of personal information that the Institute holds. Regarding the acquisition of personal information, it is premised that researchers have properly acquired it.

When researchers need personal information for their research, they shall do so by taking proper procedures, including informed consent and other necessary permission, referred to in the Ethical Guidelines and Sections 12 and 13 of Article 5 of the Guidance, the Personal Information Protection Act, and other relevant laws and guidelines.

2. Definition of Terms in the Guidelines

Among the terms defined in the Ethical Guidelines[13] and the Guidance,[14] those related to the Guidelines have been redefined based on Articles 2-1., 2-2., and 2-3. The redefined terms are listed in the Appendix. As it is assumed that the following three items are not adequately explained in the Ethical Guidelines or the Guidance, more precise definition are given to them here.

1) Handling of information on the deceased and the unborn
2) Pseudonymization
3) Anonymization

2-1. Handling of Information on the Deceased and the Unborn

Information on the deceased and the unborn is not regarded, in principle, as personal information in the Personal Information Protection Act,[15] but it is required to be handled in the same way as the personal information on living individuals dealt with in the Ethical Guidelines and the Guidance. In the Guidelines, the personal information referred to in the Ethical Guidelines and the information on the deceased and the unborn that can identify a specific individual shall be collectively defined as "Personal Information" hereinafter.

2-2. Pseudonymization

The replacement of information that can identify specific individuals (including the deceased and the unborn; the same shall apply hereinafter) among Personal Information defined in Article 2-1 or a personal identification code with a unique identifier shall be defined as “pseudonymization.”

---

[13] Article 1.2 of the Ethical Guidelines

[14] Article 1.2 of the Guidance

[15] It must be noted that the data on the deceased that can be linked to a living individual may be regarded as the individual’s personal information and therefore dealt as a subject of protection under the Personal Information Protection Act. Genome data of the deceased, for instance, may be considered relevant to that of their relatives.

A pseudonymized name is a code obtained by applying a keyed hash function(s) and other unrecoverable functions to the information that can identify a specific individual or individual identification codes.

Pseudonymized Personal Information, which consists of a pseudonymized name(s) and information other than personal identification data, is defined as “pseudonymized information.” A correspondence table that links personal identification codes and pseudonymized names is created separately.

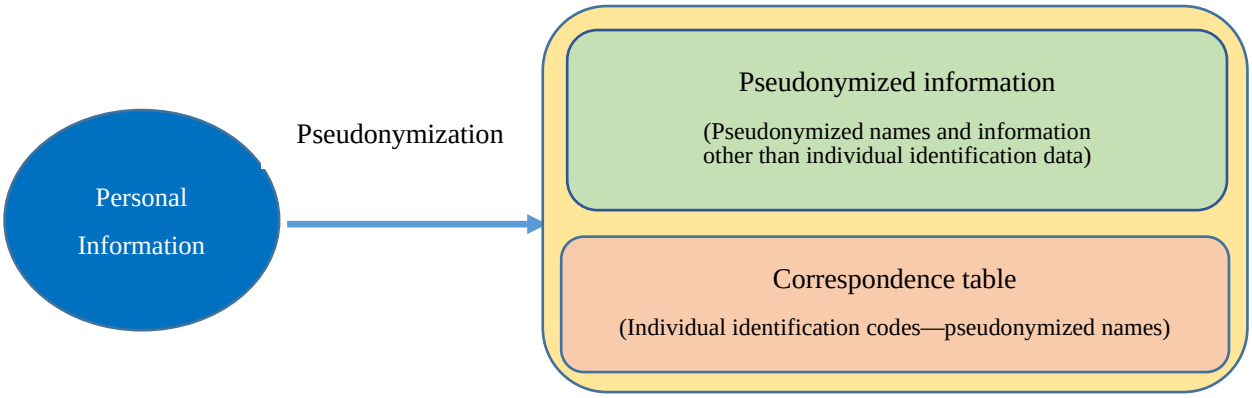


Fig. 2-2.1. Pseudonymization

A pseudonymized name is allocated to an individual in either of the following: (a) one pseudonymized name is allocated to each individual for their Personal Information; (b) more than one pseudonymized name is allocated to each individual with the elapse of time.

Generally, delinking pseudonymized information from its correspondence table makes it difficult to identify an individual in their original Personal Information from the pseudonymized information. In case (b), however, if an individual’s pseudonymized names are gathered and put together using other data (called name-based aggregation), the individual may be more likely to be identified from increased and aggregated data on the individual without a correspondence table.

2-3. Anonymization

As explained in the Ethical Guidelines and the Guidance, the process of deleting all or part of the descriptions (including personal identification codes) enabling identification of a specific individual from Personal Information (including replacing all or part of the descriptions with those irrelevant to the individual) is called “anonymization.” This includes pseudonymizing Personal Information and then, where necessary, deleting relevant correspondence tables, and where further necessary, deleting unique descriptions. This means that pseudonymization is part of anonymization.

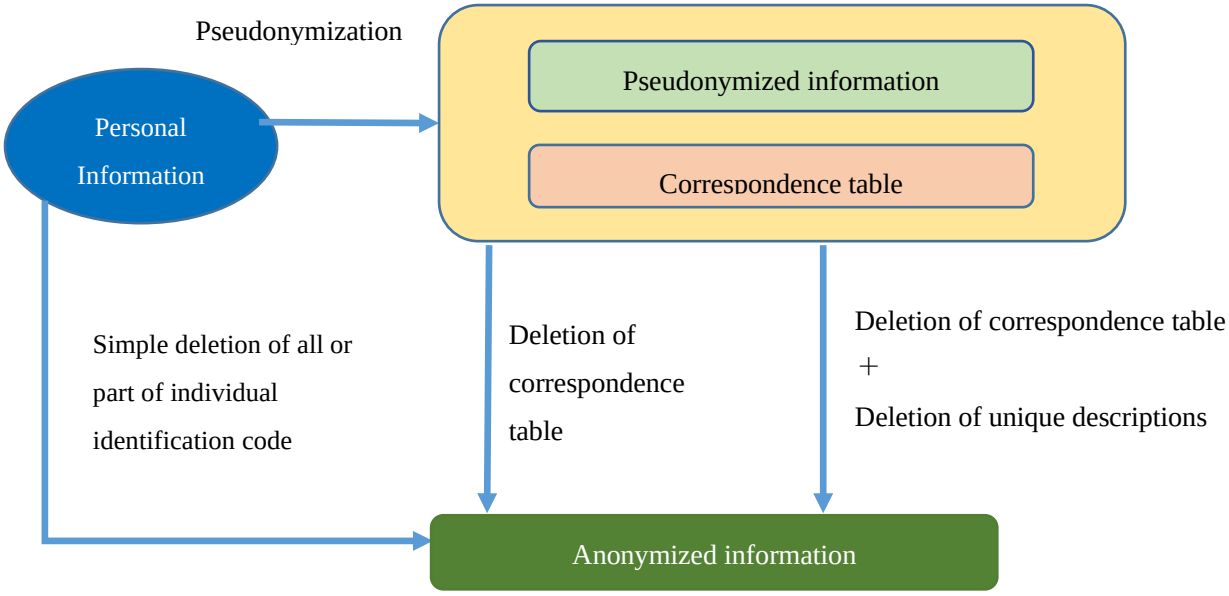


Fig. 2-3.1. Flow of Anonymization (example)

Personal Information that has been anonymized is called “anonymized information,” which includes both information that can identify an individual and information that cannot identify an individual because some anonymized information may be able to identify an specific individual by checking other data available within the Institute and codes or numbers allocated in the process of pseudonymization against a relevant correspondence table.

3. Means for Securely Managing Personal Information in Research

All Personal Information to be used for research or obtained through research shall be managed properly within an information infrastructure designated by the Institute in accordance with the Ethical Guidelines and the Guidance, and researchers shall not retain or handle such information for a private purpose.

In this article, the following topics are discussed regarding the specific means of technical security management measures[16] established by the Institute concerning the protection of Personal Information.

1) Definition of classification based on the identifiability of an individual from anonymized information
2) Security management measures based on the aforementioned classification within the information infrastructure designated by the Institute

3-1. Definition of Classification Based on the Identifiability of an Individual from Anonymized Information

The Institute shall take measures to manage retained Personal Information securely, considering the nature and degree of damage arising from leakage, etc. of the information.[17]

For this purpose, the Institute will classify Personal Information subject to anonymization according to the identifiability of an individual and take measures appropriate to each level of the classification.

3-1-1. Methods of Anonymization

Methods of anonymization include deleting descriptions that can identify a specific individual, deleting individual identification codes, pseudonymization, deleting codes that interlink data, deleting unique descriptions, and other measures based on the nature of Personal Information.[18] An

---

[16] In accordance with the Ethical Guidelines and Article 6.15.2(1) of the Guidance, the Institute must take necessary and proper measures to ensure data security. There are generally two types of security measures. One is "physical security," such as monitoring of entry (to facilities) and preventing retained Personal Information from being taken, and the other "technical security," such as control of access to Personal Information and information systems handling it, countermeasures against malware, and monitoring of information systems.

[17] Careful consideration must be given to the identifiability of an individual (i.e. the level of anonymization), the presence or absence of special care-required personal information, the nature and degree of damage due to leakage, etc. of Personal Information (the *Measures to Ensure Proper Management of Personal Information Held by Incorporated Administrative Agencies, etc.*, (Notice No. 85, issued on September 14, 2004 and Notice No. 143, partially amended on October 22, 2018 by Director-General of the Administrative Management Bureau, the Ministry of Internal Affairs and Communications)).

[18] These measures are outlined in Article 2 (Definition of Terms) of the Guidelines. Other references include the

example of the flow of anonymization is shown in Fig. 2-3.1.

3-1-2. Classification Based on the Identifiability of an Individual from Anonymized Information

The Institute classifies the original (or unprocessed) Personal Information into four levels and anonymizes it based on the classification considering the purpose of information analysis and management. The four levels (or five levels including the original Personal Information), as shown below, are determined based on the level of identifiability of a specific individual from the Personal Information that has been anonymized (or anonymized personal information).

Personal identification level 1

Pseudonymize all the single pieces of information, such as name and facial image, from each of which a specific individual can be easily identified and individual personal identification codes (pseudonymization (1)) and create pseudonymized information 1 and correspondence table 1.[19] The degree of identifiability for pseudonymized information 1 is defined as personal identification level 1.

Personal identification level 2

Delete correspondence table 1[20] and pseudonymize all the single pieces of information **other than the correspondence table** from a combination of which a specific individual can be identified, such as medical record ID, date of birth, address, and organization that the individual belongs to[21] (pseudonymization (2)), and create pseudonymized information 2 and correspondence table 2. The degree of identifiability for pseudonymized information 2 is defined as personal identification level 2.

Personal identification level 3

Delete correspondence table 2,[22] delete unique descriptions, and take other measures based on the nature of the Personal Information. The degree of identifiability after applying these means of anonymization is defined as personal identification level 3.

Personal identification level 4

Ethical Guidelines, the Guidance, and the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc. —Concerning Anonymized Personal Information Held by Incorporated Administrative Agencies, etc.* (issued by the Personal Information Protection Commission in March 2017).

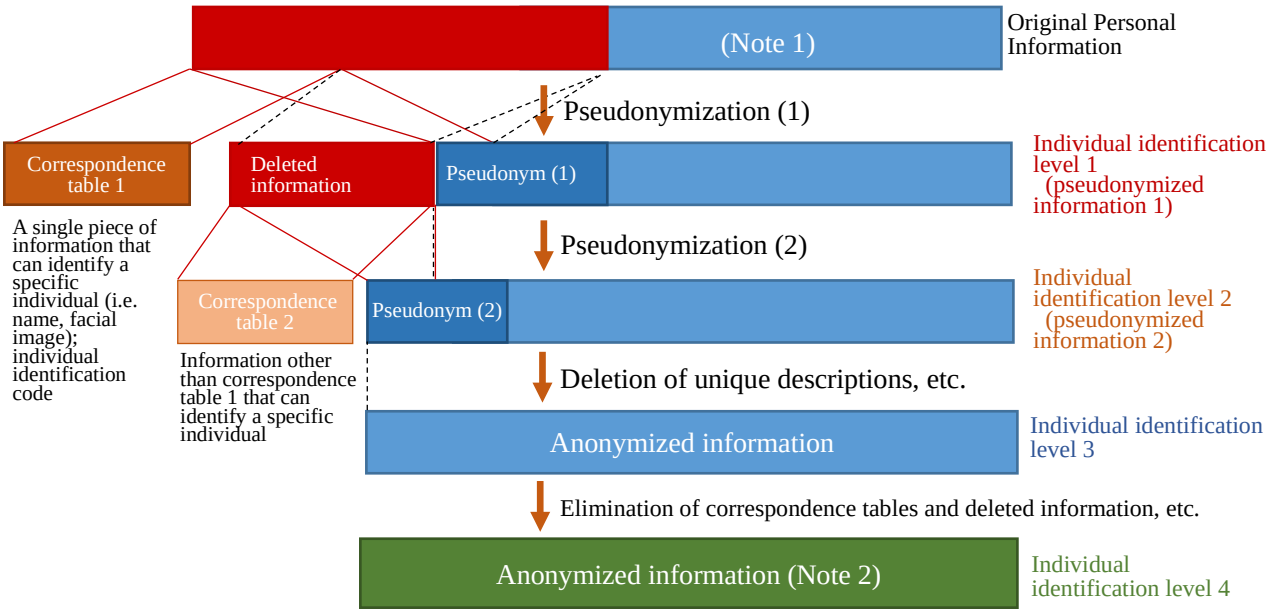
[19] At level 1, pseudonymously processed information 1 is not delinked from correspondence table 1 as pseudonymously processed information and deleted information are linked to each other through relevant pseudonymized names and correspondence table 1.

[20] When a keyed hash function is used, existing keys or the information linked to them need to be deleted. Correspondence table 1 must not be deleted; it must be kept in a separate location.

[21] For more details, see the identification codes specified in the Health Insurance Portability and Accountability Act of 1996 (HIPAA) in the United States.

[22] When using a keyed hash function, delete existing keys or the information linked to them. Note that correspondence table 1 removed from pseudonymously processed information 1 must not be eliminated; it must be kept in a separate location.

Completely erase correspondence tables and other deleted information instead of keeping them. The degree of identifiability after applying these means of anonymization is defined as personal identification level 4. Anonymized personal information (see the Appendix) belongs to this level.



Notes:
1. Original Personal Information in the blue box includes medical records, endoscopic images, X-ray images, and DNA data.
2. Limited to information that cannot identify a specific individual

Fig. 3-1.2. Interrelation Between Individual Identification Levels

In general, information at personal identification levels 1, 2 and 3 above corresponds to “anonymized information (limited to information processed or managed in such a way that the source of a sample or data cannot be determined immediately),” specified in the Ethical Guidelines and the Guidance, and information at personal identification level 4 corresponds to “anonymized information (limited to information that cannot identify a specific individual)”.

To be noted is that anonymized information, such as correspondence tables and deleted information, which is kept within the Institute is not regarded as anonymized personal information (see the Appendix) because it can be checked against other data within the Institute that can identify a specific individual.[23]

3-2. Security Management Measures for the Identification Level of the Individual

In principle, Digitized Personal Information shall physically be kept in a database or storage of a

[23] See the illustration of the table in Article 1.2(25) of the Guidance.

proper management area within the Institute.

Research data held by the Institute is classified based on the level of confidentiality as shown in Fig. 3-2.1. Researchers shall properly handle Personal Information to be used for research as specified in the information security procedure set forth by the Institute. The Personal Information held by the Institute is grouped into four categories: confidential, within division only, within the Institute only, and sharable with external parties. Except for extremely private Personal Information, which shall be dealt with as highly confidential, Personal Information may be either shared only internally (within a division(s) or the Institute), or externally (with limited external parties or within a limited time frame). Information to be shared among any people belongs to “Public.”

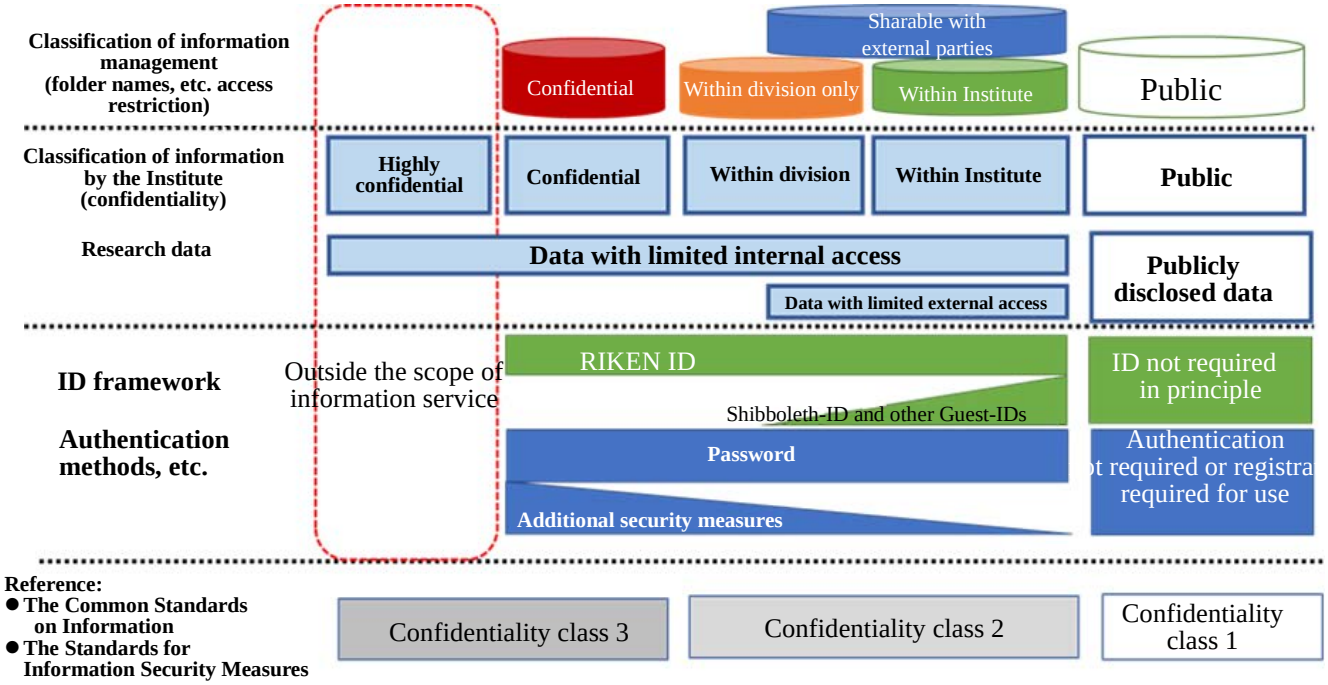
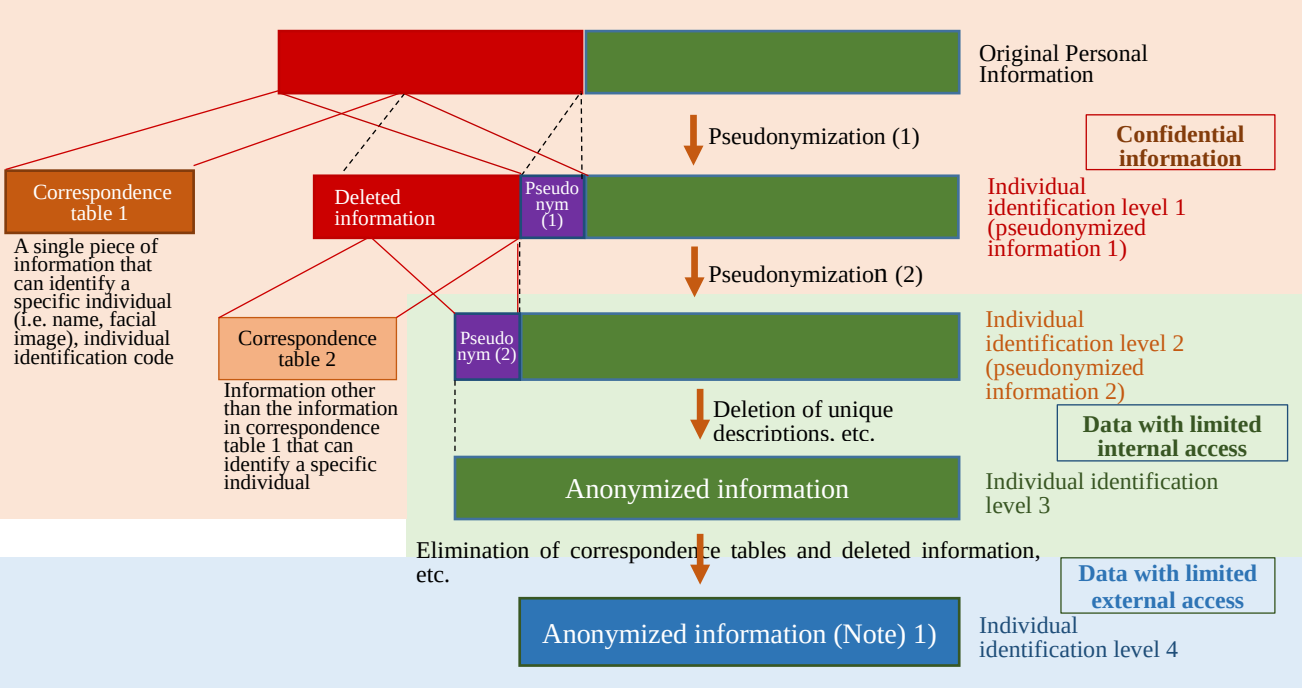


Fig. 3-2.1. The Institute’s Classification of Information Based on Confidentiality Level

3-2-1. The Institute’s Classification of Personal Information in Terms of Technical Security Management



Note:
1. Limited to information that cannot identify a specific individual

Fig. 3-2-1.1. Relationship Between Level of Individual Identification and Classification of Information Based on Security Level

As shown in Fig. 3-2-1.1., the original Personal Information (including correspondence table 1) and pseudonymized information 1 (including correspondence table 2) at personal identification level 1 shall be dealt with as confidential, and maintenance of original copies of Personal Information and pseudonymization (1) and (2) shall be carried out in the management area designated as confidentiality class 3 based on Chapter 4 of the *Standards for Information Security* (established in April 2019 by the Information Security Committee).

Pseudonymized information 2 at personal identification level 2 (including deleted unique descriptions) and anonymized information at personal identification level 3, which have undergone process (2) as well as pseudonymized information at individual identification level 2 with its correspondence table 2 eliminated shall be used only within a division(s) or within the Institute. Unique data shall be processed in the management area designated for confidentiality class 2 or higher.

Information at personal identification level 4 is regarded as data with limited external access and may copied or transferred to a mobile computer to be used in the management area designated for confidentiality class 1.

The personal information administrator in charge shall determine the scope of staff members who are allowed to access Personal Information and their access rights in accordance with the Personal Information Protection Regulations (Regulation No. 6 of March 10, 2005) and the Supplementary Security Regulations for Retained Personal Information (Supplementary Regulation No. 8 of March 10, 2005), thereby ensuring that information necessary for authentication is managed thoroughly and that Personal Information is used based on the access management rules prescribed by the Institute.

3-2-2. Entrusting an External Party to Store Personal Information

In principle, Digitized Personal Information shall physically be kept in a database or storage of a proper management area within the Institute. However, if a need should arise to have Personal Information stored by an business outside the Institute like as a private cloud operator, the Institute shall satisfy the standards for the selection of institutions entrusted with external storage and the standards for the handling of information, as specified in Article 8.1.2 of the *Guidelines for the Security Management of the Medical Information System, 5th Edition* (issued on May 2017 by the Ministry of Health, Labour and Welfare)[24]. Specifically, the Institute shall comply with the standards for the selection of institutions entrusted with external storage (Section 3) set forth in Article 1, the standards for the handling of information (Section 3) in Article 2 and the procedures in Article 6 of the Guidelines.

3-2-3. Deleting Personal Information

If Personal Information needs to be deleted, the information shall be completely erased in such a way that it will be unrestorable or illegible; likewise, if a medium (including computers and servers) that contains Personal Information needs to be disposed of, the medium shall be destroyed in such a way that the information contained in the medium will be unrestorable or illegible.[25]

In addition, the research subject that retained and managed the Personal Information shall record and retain the information on the disposal of the Personal Information, including the date, content, and method of its disposal.

4. Use of Personal Information in Research

As stipulated in Article 3 of the Supplementary Ethical Regulations for Research Involving Human Subjects (Regulation No. 128 of October 1, 2003), researchers shall use Personal Information in accordance with the content of research specified in a research plan that has been approved.[26]

When a transfer of Personal Information for research takes place between the Institute and to an external institution, the institution shall ensure that an agreement has been made between the parties on proper handling of the Personal Information. It is necessary that the agreement secures a commitment that the other party does not process any of the information to identify individuals (the act of identification).

Prior to the transfer of the Personal Information, researchers shall obtain informed consent and other necessary permission from the research subject and follow procedures as required.[27]

---

[24] https://www.mhlw.go.jp/file/05-Shingikai-12601000-Seisakutoukatsukan-Sanjikanshitsu_Shakaihoshoutantou/0000166260.pdf

[25] Article 7 of the Personal Information Protection Act; Article 5.7 of the *Measures to Ensure Proper Management of Personal Information Held by Incorporated Administrative Agencies, etc.*, (Notice No. 85, issued on September 14, 2004 and Notice No. 143, amended on October 22, 2018 by Director-General of the Administrative Management Bureau, the Ministry of Internal Affairs and Communications)

[26] See the Ethical Guidelines and Article 3.8 of the Guidance.

[27] See the Ethical Guidelines and Article 5.12 and 5.13 of the Guidance.

5. Secure Management of Personal Information for Research Provided by an External Institution

When the Institute retains Personal Information provided for a research purpose by an external institution, the Institute shall examine the content of the information to handle it properly in accordance with Article 3 of the Guidelines irrespective of how it had been dealt with by the external institution.

5-1. Procedures for Receiving Personal Information from an External Institution

Researchers may receive Personal Information from an external institution in the following procedure in accordance with the Regulations for Tangible Property Derived from Research (Regulation No. 10 of March 9, 2006) and the Personal Information Protection Regulations (Regulation No. 6 of March 10, 2005) formulated by the Institute.

(1) The Industry Partnership Division , Research Support Division, or other relevant division shall conclude a tangible property transfer agreement (or material transfer agreement; hereinafter "MTA") or confidentiality agreement in accordance with the Regulations for Tangible Property Derived from Research (Regulation No. 10 of March 9, 2006) and the Provision and Receipt of Tangible Property Derived from Research (Notice No. 8 of March 9, 2006), formulated by the Institute.

(2) When the Institute receives Personal Data based on an MTA or confidentiality agreement, the personal information administrator in charge shall ensure that a "data transfer slip" is exchanged between the parties in order to record the activities conducted in relation to the data transfer. The data transfer slip shall include the content of the data, purpose of use, names and contact information of the administrators in charge, and date and time of receipt of data.

(3) When a personal information administrator at a center of the Institute is in charge of the administration of the transfer, the administrator shall report to the Director of the General Affairs Division on the possession of personal information files in accordance with the Personal Information Protection Regulations (Regulation No. 6 of March 10, 2005).

(4) The retained Personal Information shall be classified and access rights be assigned as appropriate based on the classification of information set forth in Article 3 of the Guidelines.

Anticipating the risk of the portable device containing data being stolen or lost, researchers shall ensure that designated USB flash drives contain passwords, and that encrypted portable hard disks and tapes, or other high-security media are used in principle. The data shall be encrypted in such a way that, if the data should be lost or stolen, it will be illegible.

If there is a need to receive data through a network or other medium, however, researchers shall consult the technical compatibility of the means of receipt with the General Information Security Section (the Cybersecurity Section of the Information System Division) wherever necessary and receive Personal Information under the support of the General Information Security Section.

6. Providing Personal Information for Research to an External Institution

Researchers shall conduct research in compliance with the Personal Information Protection Regulations (Regulation No. 6 of March 10, 2005) and they may not provide Personal Information to an external institution for purposes other than the intended purpose of use.

However, if a need arises to provide Personal Information to a research institution, researchers shall follow the procedure below in accordance with the Regulations for Tangible Property Derived from Research (Regulation No. 10 of March 9, 2006) and Article 19 of the Personal Information Protection Regulations (Regulation No. 6 of March 10, 2005), formulated by the Institute.

(1) The Industry Partnership Division , Research Support Division, or other relevant division shall conclude an MTA or confidentiality agreement based on the Regulations for Tangible Property Derived from Research (Regulation No. 10 of March 9, 2006) and the Provision and Receipt of Tangible Property Derived from Research (Notice No. 8 of March 9, 2006), formulated by the Institute.
(2) When the Institute receives Personal Data based on an MTA or confidentiality agreement, the personal information administrator in charge shall ensure that a “data transfer slip” is exchanged with the external institution in order to record the activities conducted between the parties in relation to the data transfer. The data transfer slip shall include the purpose of use at the external institution, names and contact information of the administrators in charge, and the date and time of receipt of data shall be in accordance with Article 20 of the Personal Information Protection Regulations (Regulation No. 6 of March 10, 2005).
(3) Regarding technical compatibility of the means for providing Personal Information based on the MTA or confidential agreement, researchers shall consult the Director of the General Information Security Section whenever deemed necessary, and provide Personal Information under the support of the section.

Appendix. Definition of Terms

The terms that appear in the Guidelines are defined, in principle, based on the Ethical Guidelines and the Guidance. Among the terms defined in the Ethical Guidelines and the Guidance, those relevant to the Guidelines are redefined based on Article 2 of the Guidelines as follows. For details, please also refer to the Ethical Guidelines, the Guidance, or other relevant laws and regulations.

- **"Medical and health research involving human subjects"** and **"Research"** mean an activity aimed to maintain and improve the health of the people, to cure those who are injured or ill, or to acquire knowledge that contributes to the improvement of quality of life by understanding the causes of injuries and diseases (including the frequency, distribution, and impact of health-related symptoms) and clinical conditions, preventing injuries and diseases, and improving or verifying the effectiveness of diagnosis or treatment methods. "Research" in the Guidelines refers to medical and health research involving human subjects.

- **"Research subject"** means an individual (including the diseased and the unborn) as follows:
  (i) An individual that participates in research (including those who are requested to participate in research); or
  (ii) An individual whose samples or data[28] are acquired for research

- **"Research institution"** means a corporation, administrative organ, or sole proprietor that conducts research. An institution engaged partly in research-related activities, such as storing samples or data and processing data, is not regarded as a research institution.

- **"Joint research institution"** means a research institution that conducts research jointly with the Institute based on the Institute's research plans. An institution that obtains new samples or data from research subjects for such research and provides them to other research institutions is also regarded as a joint research institution.

- **"Researchers"** means those engaged in research activities (including operations by an institution engaged in gathering and providing samples and data[29]) conducted by a principal investigator. An individual outside the Institute who provides existing samples or data or is entrusted with part of the Institute's research activities is not regarded as a member of researchers.

- **"Principal Investigator"** means a researcher responsible for supervising research activities at a research institution where the researcher belongs.

---

[28] "Samples" means those obtained from the human body; "data" means information used for research.

[29] "Institution that gathers and provides samples and data" means a research institution that stores samples and data obtained from research subjects or provided by other institutions and provides them to other research institutions in a repeated manner.

- **"Personal information"** means information about a living individual that is classified as any of the following:
  (i) Information by which a specific individual can be identified based on details contained in that information, such as the name, date of birth, and other descriptive details of the individual (meaning any details, excluding individual identification codes, stated, recorded, or otherwise expressed using sound, motion, or other means in a document, drawing, or electronic or magnetic record (meaning a record kept in electronic or magnetic form (an electronic, magnetic, or any other form that cannot be perceived through the human senses alone; the same shall apply in the succeeding paragraph, item (ii)); the same shall apply hereinafter), including information that can identify an individual by comparing that information with other information);
  (ii) information containing an individual identification code.

- **"Can identify a specific individual"** means being able to specify an individual based on social convention using a piece of information or a combination of information.

- **"Can check against other data"** means being able to check information against "other data" that is held or can be obtained by the Institute using a means that the Institute considers practicable. "Other data" includes data held by other institutions, publicly known information, and information accessible by people or available at public facilities such as libraries. Data obtainable through specific investigation is not considered to belong to "other data," in general. Whether a means to identify an individual is practicable or not must be determined reasonably considering who is likely to implement the means.

- **"Individual identification code"** as used in Cabinet Order to Enforce the Act on the Protection of Personal Information (Cabinet Order No. 507 of 2003) and other relevant laws and regulations means a set of characters, letters, numbers, symbols, or other codes, as prescribed by Cabinet Order, that fall under either of the following items:
  (i) a set of characters, letters, numbers, symbols or other codes indicating a description of the physical characteristics of an individual, converted into codes in order to be processed by a computer and by which a specific individual can be identified;
  (ii) a set of characters, letters, numbers, symbols or other codes assigned in relation to the use of services or the purchase of goods by an individual, or recorded electronically or magnetically recorded on a card or document issued to an individual, each of which is uniquely assigned to each individual to enable identification of that specific user or purchaser, or of a specific recipient of the issued card or document.

<Reference>
Article 1(i) of Cabinet Order to Enforce the Act on the Protection of Personal Information (Cabinet Order No. 507 of 2003) reads as follows:
(i) Characters, letters, numbers, symbols or other codes produced by having converted any of the following physical features thereinto so as to be provided for use in computers which

conform to standards prescribed by rules of the Personal Information Protection Commission[30] as sufficient to identify a specific individual, as follows:

(a) a base sequence constituting Deoxyribonucleic Acid (alias DNA) taken from a cell;
(b) appearance decided by facial bone structure and skin color as well as the position and shape of eyes, nose, mouth or other facial elements;
(c) a linear pattern formed by an iris' surface undulation;
(d) vocal cords' vibration, glottis' closing motion as well as the shape of vocal tract and its change when uttering;
(e) bodily posture and both arms' movements, step size and other physical appearance when walking;
(f) intravenous shape decided by the junctions and endpoints of veins lying under the skin of the inner or outer surface of hands or fingers;
(g) a finger or palm print.

(omission)

(viii) Any other character, letter, number, symbol or other codes prescribed by the rules of the Personal Information Protection Commission as equivalent to each preceding item.

Regarding paragraph (viii) above, the Enforcement Rules for the Act on the Protection of Personal Information (Rules of the Personal Information Protection Commission No. 3 of October 5, 2016) has a provision prescribing health insurance certificates and the symbol and number of, and insurer's number on, an insured person's certificate and claimant certificate/beneficiary certificate

- **"Special care-required personal information"** means personal information containing statements and other particulars describing the race, creed, social status, medical history, or criminal history of an individual, or the fact that the individual is a victim of a crime, or any other statement designated by Cabinet Order as requiring special care in order to avoid exposing the individual to unfair discrimination, prejudice, or other disadvantages.

- **"Anonymized personal information**[31]**"** in the Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc. (Act No. 59 of 2003; hereinafter the "Personal Information Protection Act") means personal information (limited to personal information subject to processing) that has been processed

[30] The standards are specified in detail in the *Guidelines for the Act on the Protection of Personal Information (General)* (instituted in November 2016 and partially amended in January 2019 by the Personal Information Protection Commission)
(https://www.ppc.go.jp/files/pdf/190123_guidelines01.pdf)

[31] The term **"anonymously processed information,"** used in the Act on the Protection of Personal Information (Act No. 57 of 2003) is similar in concept to the term "anonymized personal information," used in the Personal Information Protection Act. As the Institute is subject to the provisions of the latter act, the former term does not directly apply to the Guidelines.

by a measure as specified below in accordance with its classification, so that an individual cannot be identified by that processed information and the processed information (limited to personal information subject to the provisions of the Personal Information Protection Act) cannot be restored to the state of the original personal information.

(i) Deleting a part of the descriptions or other details contained in the personal information (including replacing descriptions or other details by a means with no regularity that allows restoration of the descriptions or details to their original state); or

(ii) Deleting the whole individual identification code contained in the personal information (including replacing descriptions or other details by a means with no regularity that allows restoration of individual identification codes to their original state).

<table>
<tr><th>Type of data</th><th colspan="2">Definition</th><th>Examples</th></tr>
<tr><td rowspan="4">Personal information</td><td colspan="2">Information on living individuals that can identify a specific individual</td><td>--</td></tr>
<tr><td rowspan="3"></td><td>Identifiable via a single piece of information</td><td>Name, facial image</td></tr>
<tr><td>Identifiable in combination with other information</td><td>Information that can be checked with other information that can identify a specific individual by a correspondence table</td></tr>
<tr><td>Inclusive of individual identification code</td><td>Genome data</td></tr>
<tr><td>Special care-required personal information</td><td colspan="2">Inclusive of descriptions of which the handling of personal information requires special care</td><td>Health records, statement of medical expenses, diagnosis, genome data</td></tr>
<tr><td>Anonymized personal information</td><td colspan="2">Processed to satisfy the requirements for anonymizing personal information as specified in the Personal Information Protection Act, etc.</td><td>--</td></tr>
<tr><td>Information on the deceased and the unborn</td><td colspan="2">Not to be regarded as personal information but dealt equal to personal information in the Guidelines</td><td>(to be dealt with in the same manner as in the case of personal information and special care-required personal information）</td></tr>
</table>

Anonymized information		Deleted (including replacing descriptions with others), in whole or part, descriptions that can identify a specific person （Note: This category includes both information that can identify a specific individual and information that cannot identify a specific individual）	Pseudonymized name in place of real name
	Anonymized information （that must be processed or managed so that a research subject remains unidentified via samples or data obtained from the subject）	Anonymized information that has been processed so that a single description cannot immediately identify a specific research subject; when a correspondence table is included, it must be managed properly. （Note: This category includes both information that can identify a specific individual and information that cannot identify a specific individual）	Information generated through anonymization as described on the left

Supplemental fig. 1-1.1. Classification: personal information, anonymized personal information, and anonymized information, etc.[32]

- **"Deleting descriptions or other details that can identify a specific individual"** means deleting all or part of the descriptions or other details contained in Personal Information (including replacing descriptions or other details by a means that eliminates regularities that allow restoration of the descriptions or details to the original state[33]).

[32] The table was created based on Article 1.2 of the Guidance and the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information (General)* (issued in November 2016 (partially amended in March 2017) by the Personal Information Protection Commission) (https://www.ppc.go.jp/files/pdf/guidelines01.pdf).

[33] A temporary ID may be used only after all regularities that allow original descriptions to be restored are eliminated. When a hash function is used, for example, to generate a temporary ID from descriptions unique to an individual, such as their name, address, or telephone number, the hash function must not be used for the generation of other temporary IDs. This is because using the same hash function may yield a regularity that may retrieve original descriptions. A solution to eliminate such risk will be to use a hash function after other descriptions like random numbers are applied to the original descriptions (i.e. name, telephone number). When the same random numbers or other descriptions are applied for the hash functions to be used, it is recommended that a different combination of descriptions be provided to each business and that the combination be changed periodically in order to prevent regularities from occurring that may allow original descriptions to be restored through the random

This is an appropriate processing means for preparing anonymized personal information, which is specified in Article 10.1[34] of the Rules on the Provision of Anonymized Personal Information Held by Incorporated Administrative Agencies, etc., in accordance with the provisions of Article 44.10.1 and Chapter IV-2 of the Personal Information Protection Act.

<Practical examples of data processing>
Example 1) Process retained personal information including names, addresses, and years, months, and days of birth as follows:
1) Delete names.
2) Delete addresses or replace them with the names of prefecture and city only.
3) Delete years, months, and days of birth or replace them with years and months of birth by deleting days of birth.[35]
Example 2) Process retained personal information including names, addresses, and telephone numbers as follows:
1) Delete names and telephone numbers.
2) Delete addresses or replace them with the names of prefecture and city only.

- **"Deleting an individual identification code"** means deleting the entire individual identification code contained in Personal Information (including replacing the individual identification code with other details by a means that eliminates regularities that allow the individual identification code to be restored to its original state).

  This is an appropriate processing means for preparing anonymized personal information, which is specified in Article 10.2[36] of the Rules on the Provision of Anonymized Personal Information Held by Incorporated Administrative Agencies, etc., in accordance with the provisions of Article 44.10.1 and Chapter IV-2 of the Personal Information Protection Act.

- "Deleting a code that interlinks information" means deleting a code that interlinks Personal Information and the information that can be obtained from the Personal Information by a means (limited to codes interlinking data available for handling by incorporated administrative agencies, etc.) (including replacing the code with another that cannot interlink Personal Information and the information obtainable from the Personal Information by a means that eliminates regularities that enable the codes from being restored to its original state.

---

numbers or other descriptions. This recommendation applies in all cases where a temporary ID is used.

[34] See Article 3.2.1 of the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc. —Concerning Anonymized Personal Information Held by Incorporated Administrative Agencies, etc.* (issued in March 2017 by the Personal Information Protection Commission)
(https://www.ppc.go.jp/files/pdf/guidelines06.pdf)

[35] Rewriting original descriptions in a more general and abstract manner, as shown in the example, is a valid option.

[36] See Article 3.2.2 of the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc. —Concerning Anonymized Personal Information Held by Incorporated Administrative Agencies, etc.* (issued in March 2017 by the Personal Information Protection Commission)

This is an appropriate processing means for preparing anonymized personal information, which is specified in Article 10.3[37] of the Rules on the Provision of Anonymized Personal Information Held by Incorporated Administrative Agencies, etc., in accordance with the provisions of Article 44.10.1 and Chapter IV-2 of the Personal Information Protection Act.

<Practical examples of data processing>
Example 1) Manage basic information (i.e. name) included in personal data files separately from other information. Delete information management IDs that are used to interlink them.
Example 2) Delete the codes that are used to link the pseudonymized names used for the same individual that are aggregated from different databases.

- **"Deleting unique descriptions or other details"** means deleting unique descriptions (including replacing the descriptions with other descriptions that eliminate regularities that enable the unique descriptions to be restored to the original state).

  This is an appropriate processing means for preparing anonymized personal information, which is specified in Article 10.4[38] of the Rules on the Provision of Anonymized Personal Information Held by Incorporated Administrative Agencies, etc., in accordance with the provisions of Article 44.10.1 and Chapter IV-2 of the Personal Information Protection Act.

<Practical examples of data processing>
Example 1) Delete information particular to a given household (i.e. a family with 10 children or more).
Example 2) Replace "age 116" with "age 90 or older."
Example 3) When a specific individual can be easily identified from data, such as rare diseases, dates and times of medical examinations, hospital names, and names of the doctors in charge, delete them or process them so that the individual cannot be identified from the data.

- **"Other measures based on the nature of Personal Information"** means measures to be taken to modify to a reasonable extent the Personal Information that cannot identify a specific individual and cannot be restored to the original state. Specifically, the Personal Information may be modified, such as by deleting descriptions that can identify a specific individual, deleting individual identification codes, deleting codes that interlink data, or deleting unique descriptions.

  This is an appropriate processing means for preparing anonymized personal information, which is specified in Article 10.5[39] of the Rules on the Provision of Anonymized Personal Information Held by Incorporated

---

[37] See Article 3.2.3 of the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc. —Concerning Anonymized Personal Information Held by Incorporated Administrative Agencies, etc.* (issued in March 2017 by the Personal Information Protection Commission)

[38] See Article 3.2.4 of the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc. —Concerning Anonymized Personal Information Held by Incorporated Administrative Agencies, etc.* (issued in March 2017 by the Personal Information Protection Commission)

[39] See Article 3.2.5 of the *Guidelines for the Personal Information Protection Act on the Protection of Personal Information Held by Incorporated Administrative Agencies, etc. —Concerning Anonymized Personal Information*

Administrative Agencies, etc., in accordance with the provisions of Article 44.10.1 and Chapter IV-2 of the Personal Information Protection Act.

Specifically, the above-mentioned measures include but are not limited to the following methods:

Type of method	Description
Deletion of an item, record, or cell	Deletion of an item: deleting descriptions of Personal Information subject to processing (i.e. deleting age data entirely from personal information); deletion of a record: deleting the entire information of a specific individual; or deletion of a cell: deleting age data of a specific individual
Generalization	Replacing with a broader concept or numeric data or rounding off descriptions included in the information subject to processing (i.e. replacing "electric chief engineers classified into three classes (1st class to 3rd class) based on the accreditation system" with "electric chief engineers")
Top or bottom coding	Capping the values at the high or low end of the range of data in a personal information file subject to processing at an arbitrary set value. (i.e. lumping together groups of people of age 80 or older into one collective group of age 80 or older)
Micro-aggregation	Classifying retained personal data of a personal data file subject to processing into groups and then replacing each group with a description that represents the group
Data swap	Stochastically replacing descriptions with other descriptions that are both included in personal data files subject to processing
Addition of noise (error effects)	Replacing the original data with arbitrary data generated by adding a random number(s) conforming to a certain distribution
Generation of pseudo-data generation	Generating artificial synthetic data and including it in the personal data file subject to processing

<Practical examples of data processing>

Example 1) (Deletion of an item, record, or cell): Deleting within a designated scope an individual's movement history included in the Personal Information subject to processing when the history includes data that may help figure out the individual's home address or work address, resulting in identifying the individual or enabling their Personal Information to be restored to the original state.

Example 2) (Top coding): Replacing "the height of 170cm" of a student who is much taller than other students, for instance, with "the height of 150cm or more," when processing the personal data files including the data of body height measurements of elementary school students, in order to prevent the student from being identified or the student's Personal Information from being restored to the original state.

Held by Incorporated Administrative Agencies, etc. (issued in March 2017 by the Personal Information Protection Commission)

- **“Deleted information”** means a description or the like or an individual identification code that has been deleted from Personal Information in the process of anonymization (including replacing descriptions or other details by a means with no regularity that allows restoration of the descriptions or details to their original state).

- **“Method of processing”** means information on how Personal Information has been anonymized. Included in the methods of processing are “deleting descriptions that can identify a specific individual,” “deleting individual identification codes,” “deleting codes that interlink data,” “deleting unique descriptions,” and “other measures to be taken based on the nature of Personal Information.”

- **“Deleted information, etc.”** is an umbrella term for deleted information and methods of processing.

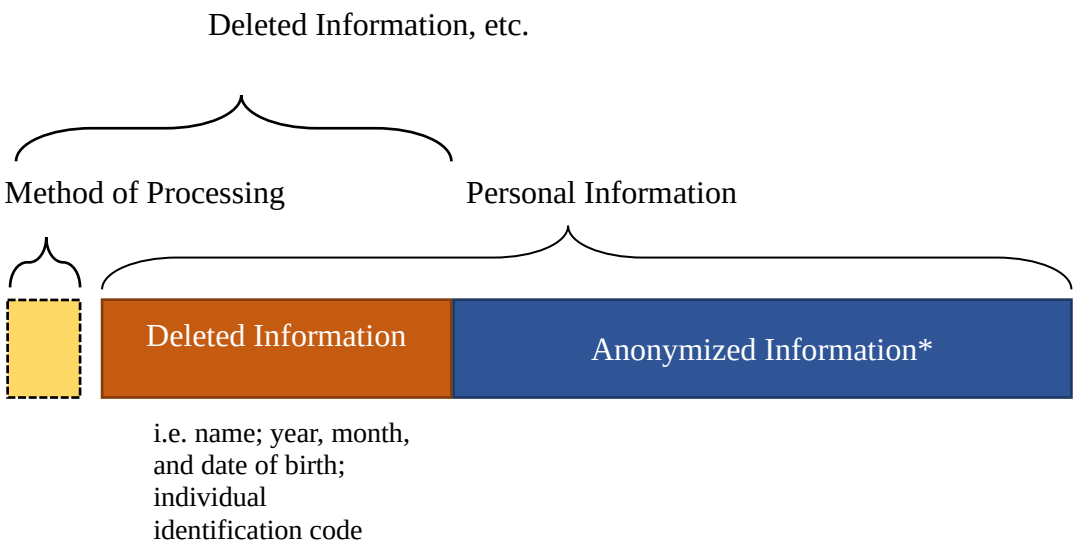


Fig. 2-2.2. Data components: Personal Information, anonymized information, deleted information, etc.

*Pseudonymized names are part of pseudonymized information, in which anonymized information may be called pseudonymized information.