DATA PRIVACY IN RESEARCH

Questions & Answers[1]

PA Sy, FM Nicolas, JM Regnim, JC Navera, & A Caraan
University of the Philippines

General

1. *Are you following GDPR?*

Yes, as a general standard of privacy practices in research involving human subjects. The General Data Protection Regulation (GDPR) is a set of guidelines for the collection and processing of personal data from individuals living in the European Union (EU). Although the Philippines' Data Privacy Act (DPA) of 2012 has been based on the Data Protection Directive (Directive 95/46/EC), GDPR's predecessor, aligning your interpretation of local privacy regulations with the GDPR is a proactive measure for the research community. Many current international privacy guidelines (of which GDPR is the de facto standard) tend to provide robust protection to data and research subjects. Nothing in the 2012 DPA conflicts with the GDPR. The EU law is only broader in scope and is *generally* consistent with many privacy laws in other countries (including the Philippines).

2. *What is a Legitimate Purpose Test?*

"Legitimate purpose"[2] refers to the processing of information "compatible with a declared and specified purpose which must not be contrary to law, morals, or public policy" (Sec 18

[1] Last updated: 5 September 2021. This is a snapshot of its 'live' version at privacyph.org/qanda. You may also have to see its primary documents: the Primer and the Toolkit. A self-paced introductory course on Data Privacy in Research is also available at privacyph.org/course. Related in-person or online workshops (1-day, 2-day, or some other options) are available upon request. Clarificatory, lingering questions arising from these materials are addressed here. See privacyph.org for more updates.

Acknowledgments: Shiela Lalaguna (layout) and Crizza Elaine Ilustre (graphic illustration).

Disclaimer: Readers are advised to treat the foregoing discussions as offered professional opinions formulated with the best evidence and references available to their authors. The National Privacy Commission (NPC) and the Supreme Court remain the "final" legal authorities on the subject of data privacy.

[2] In the language of GDPR, "legitimate interests". The processing of personal information is legitimate if it "is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child" (Article 6 (1)(f), GDPR).

b, Implementing Rules and Regulations (IRR) of the Data Privacy Act of 2012). By subjecting a project, program, or system to the Legitimate Purpose Test one can determine whether its purpose or interest is legitimate or not. The National Privacy Commission (NPC) stipulates that such a test (Figure 1) is a totality of 3 sub-tests (purpose, necessity, and balance).[3]

Figure 1. The questionnaire for the Legitimacy Purpose Test

Purpose Test

	Answer
What interest or purpose does your project, program, system serve?	
What does the processing of personal information seek to achieve in relation to the declared and specified purpose?	

Necessity Test

	Yes	No	Remarks
Is the processing of personal information *necessary* for the purpose or the interest being pursued?			
Besides processing personal information, is there any other way to achieve the identified interest or declared purpose?			

Balance Test

	Yes	No	Remarks
Does the interest of the project, program, system override the rights and freedom of the data subjects?			
Other factors that may unfairly impact the data subject			

[3] National Privacy Commission, Advisory Opinion 2020-050. https://www.privacy.gov.ph/wp-content/uploads/2020/12/Redacted-Advisory-Opinion-No.-2020-050.pdf. See also: University College London. (2019, April 3). Legitimate interest as a lawful basis. Data Protection. https://www.ucl.ac.uk/data-protection/guidance-staff-students-and-researchers/practical-data-protection-guidance-notices/legitimate

Does the nature of the interest or purpose affect the data subject negatively?			
Can the processing (collection, use and retention) of personal information cause unwarranted or severe harm to the data subject?			
Are there privacy and security safeguards[4] put in place?			

Passing purpose, necessity, and balance tests is passing the Legitimate Purpose Test. You should run this Test with your data protection officer (DPO) or her team.

Given its privileged position,[5] research could take the presumption of legitimacy and regularity[6] when "*intended* for a public benefit"[7] and subjected "to the requirements of applicable laws, regulations, or *ethical standards*" (Sec 5c, IRR; emphasis added). However, subjecting data processing activities (even in or for research) to the Legitimate Purpose Test is an exercise in prudence.[8]

3. What is the fiduciary duty of personal information controllers (PIC) to data subjects?

Personal data processing entails entering into an ethical and legal relationship. At *least* 2 parties are involved: the personal information controller (PIC) - first party (information

[4] Such safeguards may include data minimisation, de-identification, encryption, hashing, data retention limits, access restrictions, opt-out options. Without such safeguards data subjects could be unfairly exposed to privacy and security risks.
[5] In GDPR, for instance, an organization *may* be permitted to process personal data for research purposes without the data subject's consent (Article 6(1)(f); Recitals 47, 157).
[6] *Omnia praesumuntur rite et solemniter esse acta donec probetur in contrarium*—the "presumption of regularity" principle maintains that, unless there is evidence to prove the contrary, transactions in the course of business are assumed to have been conducted in the usual manner (See, for example, *People v. De Guzman*, G.R. No. 106025, February 9, 1994, 299 *SCRA* 795, 799). Research is the business, the official duty of researchers (cf. Rep. Act No. 8439). The "public benefit" and "ethical standards" (operationally meaning, the conduct of ethics review) qualifiers *may* exempt certain research projects from having to go over the Legitimate Purpose Test. Considerations of purpose, necessity, and balance are part of "the usual manner" of research ethics review. Hence, the combined presumption of legitimacy and regularity. In research, this presumption is necessary especially in certain contexts where it is "often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection" (Recital 33 of the preamble of the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC).
[7] The Philippines has yet to develop a robust "public benefit test," especially for data privacy and research. The Revenue Regulations 3-98 (Dec 8, 1998), however, identifies certain pursuits as requirements for organizations to acquire public benefit status. Consistent with the stipulation are regulations in many countries on NGO activities as "public benefits", including relief of poverty, advancement of education. advancing health and saving lives, promoting community interests, and so on. Aligning research with the UN's Sustainable Development Goals (SDG) is *prima facie* beneficial to the public. In short, a researcher cannot just arbitrarily self-declare a research project as one of "public benefit" without any accountability.
[8] See the prudent-person rule under Question 3 below.

fiduciary) and the data subject - second party (beneficiary). For purposes of research and other personal information processing activities, the data subject or human participant effectively entrusts his personal data to the researcher or PIC.[9]

In research involving human participants, beneficence can be direct or indirect.[10] But even setting aside beneficence for a moment, as information fiduciary, the PIC or the researcher is bound by a set of duties aimed at ensuring that all the data processing under his watch is fair as well as in the best interests of the beneficiary and in good faith.[11]

The duties of information fiduciaries are: (i) duty of confidentiality, (ii) duty of loyalty, and (iii) duty of care.[12] The **duty of loyalty** is the most basic of fiduciary duties, requiring the fiduciary to put the interests of the beneficiary first, ahead of his own and refraining from exploiting the relationship for his personal benefits.[13] The duty of loyalty entails that the researcher acts in the best interests of the data subject or research participant and design the research project, program or system in such a way that avoids creating conflicts of interest with the beneficiary.

As an information fiduciary, the researcher also owes a **duty of care** in carrying out research. In any fiduciary relationship, the fiduciary is expected to at least exercise the same amount of care that any *prudent person*[14] would exercise in a similar position or under similar circumstances. In the Philippine Civil Code, where "the law or contract does not state the diligence or standard of care to be observed in the performance of [a fiduciary] obligation, the expected norm of action to be required is that "a good father of the family."[15]

For researchers, the duty of care is also reflected in the ethics review process conducted by research ethics committees (REC). The review, in part, seeks to ensure that the researcher or the research institution is discharging the duty of care to data subjects or research participants in line with their rights as well as with the principles prescribed in local and international codes of conduct.[16]

---

[9] That personal data can be considered as property, is a concept that can be traced back to John Locke's theory of property. Every person has a property in his *own person* and therefore has inalienable, fundamental human rights (Locke, *Two Treatises of Government*, II.v.27). Contemporary privacy regulations, however, protect personal data not simply as private property but as part of the data subject's "inviolate personality" (Warren & Brandeis, The Right to Privacy, 1890) irreducible to material consequences of property transactions. Data privacy is a legal right.

[10] In clinical research, the fiduciary obligations of physician-researcher to patient-subject are especially strong. See, for instance, Miller, P. B., & Weijer, C. (2006). Fiduciary obligation in clinical research. *The Journal of Law, Medicine & Ethics: A Journal of the American Society of Law, Medicine & Ethics*, *34*(2), 424–440. https://doi.org/10.1111/j.1748-720X.2006.00049.x

[11] Valsan, R, "Fiduciary Duties" in Marciano, A. (2019). *Encyclopedia of law and economics*. Springer Berlin Heidelberg.

[12] Balkin, J. M. (2020). The Fiduciary Model of Privacy. *Harv. L. Rev. F., 134*, 11.

[13] Mariani, J. F., Kammerer, C. W., & Guffey-Landers, N. (2010). Understanding Fiduciary Duty. *Florida Bar Journal*, 84, 20-21.

[14] Analogously in business, the Prudent-Person Rule (aka "prudent investor rule") restricts the choices of the financial manager (fiduciary) of an owner's or beneficiary's account to certain types of investments or preferences set by the trust, avoiding reckless speculations (Law, J. (Ed.). (2016). *A dictionary of business and management*. Sixth ed. Oxford University Press).

[15] Abrogar v. Cosmos Bottling Company, G.R. No. 164749 (2017). See also Art. 1173 of the Civil Code.

[16] Slowther, A., Boynton, P., & Shaw, S. (2006). Research governance: ethical issues. *Journal of the Royal Society of Medicine, 99*(2), 65-72.

The Data Privacy Act safeguards *all* personal information. The leeway the law provides for research does not mean any blanket exemption for research, with no consideration for the rights of research participants. The researcher has a **duty of confidentiality**, and research ethics committees are tasked in part with ensuring that the basic standards of data privacy are complied with.

Confidentiality is a fundamental principle in healthcare as well as in research. Without it, the patient-doctor relationship cannot be built on trust; health information essential to the treatment of medical conditions cannot be entrusted without fearing privacy breaches. Without confidentiality, research subjects might not feel comfortable sharing personal information to researchers. Hence, keeping such information in confidence is an important obligation of both healthcare workers and researchers.[17]

Sometimes tensions arise between the researcher's duty of confidentiality and the very research process that's ultimately about sharing knowledge. Even in healthcare, certain overriding medical and legal obligations (for instance, in cases involving reportable diseases, court orders, child abuse) may undermine the duty of confidentiality. In both research and healthcare, the information fiduciary's duty of confidentiality is irreducible to any absolute prohibition against sharing personal data. Rather, what the duty of confidentiality ensures is that information would only be shared with specific persons for specific legitimate purposes.[18] (See Legitimate Purpose Test in Question 2 above.)

Finally, the information fiduciary duties of loyalty, care, and confidentiality have serious ramifications for the researcher's handling of the *informed consent process* as well as for the overall legitimacy of personal information processing in research. While it may be argued that informed consent (as a document to be signed) is not always strictly required, especially in research that does not require the active, direct involvement of data subjects (see discussion on invisible processing in **Question 5** below), the "default position" is to insist on informed consent.[19] The legitimacy of information processing and the assumption of fiduciary duties still rest on it as a "freely given, specific, informed indication of will, whereby the data subject agrees to the collection and processing of personal information about and/or relating to him or her."[20] The data subject's valid consent gives the researcher a direct authority to exercise specified discretionary powers concerning the data subject's personal information. By providing consent, the data subjects are investing their *trust* in research, establishing a fiduciary *relationship* between the beneficiary (data subject) and the fiduciary (the PIC or the researcher).

---

[17] World Health Organization. (2009). *Research ethics committees: basic concepts for capacity-building*. World Health Organization.
[18] Williams, G., & Pigeot, I. (2017). Consent and confidentiality in the light of recent demands for data sharing: Consent, confidentiality, and data sharing. *Biometrical Journal, 59*(2), 240–250. https://doi.org/10.1002/bimj.201500044
[19] Ibid.
[20] Sec. 3(b), R.A. 10173

Researchers

4. *Am I required to conduct a Privacy Impact Assessment (PIA) for my research project?*

No, especially *if* your research project involves less than 1,000 research participants,[21] uses their personal data for the sole purpose of completing the project for which you obtained their consent, and makes no decisions directly affecting them. However, your research organization is *expected* to run periodic PIA for its data collection and utilization practices that may include ***your own*** research project.

The PIA is an internal organizational exercise in due diligence and risk management. Your organization may set its own requirements for proper appreciation of your own data privacy situation.

However, not conducting a formal PIA does not mean dispensing with appropriate privacy and security measures. Privacy By Design[22] applies to all data processing activities.

For activities that are likely to trigger PIAs, see **Question 9** below.

5. *Does 'invisible processing' violate data privacy rights?*

Yes. "Invisible processing" refers to any act of gathering, using, or processing personal data without the knowledge of the data subject. Absent any compelling, legitimate reason (Legitimate Purpose Test; see **Question 2** above), *not* obtaining the data subject's consent (and, for research, ethics approval), invisible processing is *prima facie* a violation of data privacy. Uninformed, data subjects cannot assert their rights.[23]

The 'gold standard' modality for data processing necessitates the engagement of data subjects through the informed consent process. Other than consent, however, there are legitimate bases for collecting and processing personal information--including contracts, legal obligations, vital interests, or any public duty necessitating the processing of personal data. But without consent, the onus rests solely on the personal information controller (PIC) passing not only the Legitimate Purpose Test but also keeping his Fiduciary Duty (see

[21] 1,000 individuals (or 250 employees) is the threshold set by the National Privacy Commission (NPC) to require a personal data-processing system to be registered with the Commission. We're using the same figure similarly here as a rule of thumb to distinguish "small" from "big" research projects that might need a formal PIA.
[22] See Privacy By Design In Research: privacyph.org/pbdresearch
[23] Information Commissioner's Office. (n.d). When do we need to do a DPA. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/when-do-we-need-to-do-a-dpia/#when10 (accessed 22 July 2021).

Question 3 above) to the data subject, *at least* not doing anything that runs counter to the data subject's interest.

If you legitimately brought personal data from other organizations or obtained datasets through publicly accessible sources, to the extent feasible, try to contact the data subjects concerned and provide them with your organization's privacy notice. If not feasible or if obtaining consent would entail disproportionate efforts, you must *at least* conduct a PIA and ensure that privacy risks are mitigated.

Most of all, have your research proposal reviewed and approved by a competent research ethics committee. The ethics review process, in part, seeks to ensure *fairness* in processing of personal data, using nondiscriminatory research methodology and fair selection of human or data subjects.

6. How long can personal data be stored for research purposes?

In general, completed study-related documents are "archived for a minimum of three years."[24] So that's a reasonable period of keeping personal data in research, too, provided appropriate privacy and security measures are put in place. For healthcare purposes, the Health Privacy Code[25] stipulates that all medical records, whether in electronic and/or paper format, shall be stored for fifteen (15) years. For personal data used in medico-legal cases, records shall be stored for a lifetime. So, for research using healthcare data, the archival requirements of healthcare take precedence over research.

However, once personal data obtained from research are properly *de-identified* using appropriate technical means, theoretically, they can be kept forever, shared widely and used for further research.

7. Am I allowed to break confidentiality?

Yes, and *only* under special circumstances where it can be *shown* that doing so is in the best interest of the subject or the public may a researcher be allowed to *legally* break confidentiality. These circumstances may involve (but not limited to) risks of serious harm to the subject or others. In the Philippines, researchers have been advised to disclose to subjects the "legal or other limits to the researcher's ability to safeguard confidentiality,

[24] Sec 14.2.1 of the *National Ethical Guidelines for Health and Health-Related Research* (2017).
[25] DOH, DOST, & PhilHealth (2016). Joint Administrative Order No. 2016-0002. http://ehealth.doh.gov.ph/images/HealthPrivacyCode.pdf

and the possible consequences of breaches of confidentiality."[26] Legally[27] breaking the information fiduciary's duty of confidentiality (see **Question 3** above), however, does not always mean direct, personal intervention by the researcher on the subject or her condition. It can mean reporting the matter to authorities or referring to more appropriate professional or institutional support, while maintaining prudence in the pursuit of research.

Privacy Impact Assessors

8. *In the use of the Toolkit, particularly Sec 4.1.1, do we have to include or indicate all the personal data in our possession?*

No, it is not practical to do so. The expectation is to *take stock* of the personal data in your inventory, noting general and special characteristics, volume, types, examples, and so on. Identities that might be of special interest in the inventory include prominent individuals or those with elevated risk of being targeted for fraud. Hence, the inclusion of specific names or *data types* in the sample form.

9. *What activities or projects are likely to trigger PIAs?*

For data processing that can potentially harm individuals, Privacy Impact Assessment should be considered. A PIA is one demonstration of how research organizations, or any institution for that matter, incorporates data privacy protection throughout the data life cycle and into their programs, projects, or systems. Generally, PIA is recommended when large-scale personal data collection is involved, but such assessment does not stop in the collection phase. Significant changes in technologies used or in organizations tasked with the use, storage, and disclosure of data are also considered.

For data holdings, PIA is recommended when conversion of data from physical to electronic (and vice versa) is undertaken. Processing activities such as linking or reverting of pseudonymized data into an identifiable form, or adding new types of identifiable information to a previously anonymized database can trigger an assessment.

---

[26] *National Ethical Guidelines for Health and Health-Related Research* (2017), Sec 12.12.

[27] The use of the qualifier "legally" here is deliberate. In practice, however, researchers are much more conflicted in their beliefs than what we can readily discuss here. Their ethical positions do not always align with what's legally permissible. See, for instance, Surmiak, A. (2020). Should we Maintain or Break Confidentiality? The Choices Made by Social Researchers in the Context of Law Violation and Harm. *Journal of Academic Ethics, 18*(3), 229–247. https://doi.org/10.1007/s10805-019-09336-2. Moreover, the 'ethics-first' doctrine of strict confidentiality tends to oppose legal justifications for breaking the researcher's duty (Lowman, J., & Palys, T. (2014). The betrayal of research confidentiality in British sociology. *Research Ethics, 10*(2), 97–118. https://doi.org/10.1177/1747016113481145).

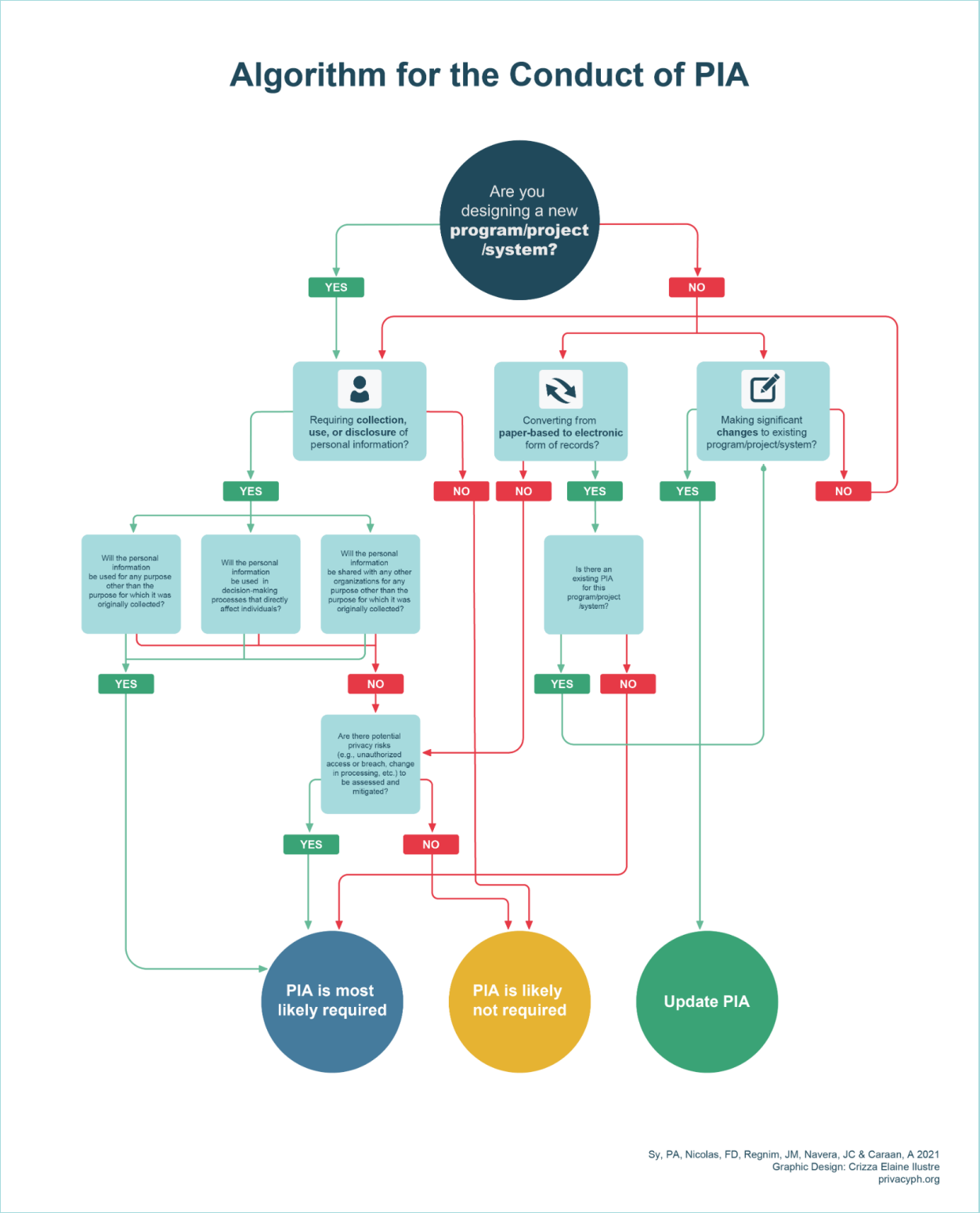
You may consider conducting a PIA if databases with identifiable personal data are merged. Databases obtained from multiple commercial or public sources can likely result in greater identifiability of individuals, thus increasing privacy risks.

For information systems, PIA can be conducted to reflect significant developments to or applications of personal-data processing technologies that provide greater data accessibility. These can result in changes in the collection, use, and sharing of identifiable personal information among researchers or organizations.

Below is an algorithm for deciding whether an activity, program, project, or system involving personal data would require a PIA.

Figure 2. Algorithm for deciding whether the conduct of PIA is warranted.[28]

[28] A higher resolution image of this illustration is at privacyph.org/pia-algo.png



Whether already known as maximum (high) risk or not, these activities tend to trigger PIA.

10. What activities are considered as maximum (or high) risk?

Maximum (or high) risk activities can lead to data breaches or privacy violations, and the GDPR stipulates that these activities are likely to require the conduct of a PIA.[29] Such activities include the automated processing of personal data, large-scale profiling or processing of sensitive personal data (e.g., race, religious beliefs, biometrics,medical records, criminal records), and large-scale monitoring of publicly accessible areas. Low (negligible) risk processing can at times turn to high (maximum) risk, for instance, in cases involving changes in technologies for user authentication, applications for system management, and merging or matching of personal information or identifiers on a grand scale from multiple sources.

While systematic profiling does not necessarily lead to the conduct of PIA, you must also consider whether the collection or information processing is extensive, includes sensitive personal information, or poses significant risks to data subjects. Such risks can harm a subject's reputation, health, finance, safety, and the like.

To elaborate on risky activities, refer to the table below (Figure 3).

Figure 3. Processing Operations Considered Maximum Risk[30]

Type of Processing Operation	Areas of Application
Processing of Biometric Data, Genetic Data	Thumb marks or fingerprints, face recognition system for attendance or surveillance purposes Medical records of a patient for diagnostic and treatment purposes
Processing of Personal Data from Vulnerable Groups	Children who are victims of sexual exploitation remaining anonymous on the news HIV or AIDS carriers' medical history[31]
Automated / Machine-based Processing	Scraping profiles from social media Credit history and purchase records

[29] GDPR Art. 35 (3).
[30] Cf. Information Commissioner's Office. (n.d.) Examples of processing 'likely to result in high risk'. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/examples-of-processing-likely-to-result-in-high-risk/ (accessed 31 July 2021).
[31] See Art 6, Section 44(a), Rep. Act No. 11166

	QR-codes for contact tracing or admission in establishments[32]
Large-scale Processing	National elections Social media data used for targeted ads and campaigns;[33] ads that are served based on one's online activities Public wifi in public spaces such as libraries, parks and hotels
Invisible or Indirect Processing	Collection of personal data from publicly available sources Aggregators gathering data from cookies and selling them to a market not disclosed to individuals concerned
Data Matching	Requirement and storage of 2 identification documents for identity validation[34] CCTV surveillance and dashcam footage for law enforcement or identification of suspects
Tracking	GPS trackers of mobile applications (Waze, Facebook, fitness monitors) for the provision of service or market Cookies to keep records of the browsing history of an individual.

[32] See, for instance, Kabagani, L.J. "Just one app needed in Mandaluyong, Pasig, Valenzuela, Antipolo". *Philippine News Agency*, March 4, 2021, https://www.pna.gov.ph/articles/1132624 (accessed 31 July 2021).
[33] See, Morales, N. J., "Philippines' watchdog probes Facebook over Cambridge Analytica data breach". Edited by Petty, M. and Meijer, E, *Reuters*, April 13, 2018, https://www.reuters.com/article/us-facebook-privacy-philippines-idUSKBN1HK0QC (accessed 14 August 2021)
[34] See, for instance, Agna, K. S. "ARTA, key agencies to improve data-matching system among gov't agencies". Anti-Red Tape Authority, July 20, 2021 https://pia.gov.ph/press-releases/2021/07/20/arta-key-agencies-to-improve-data-matching-system-among-govt-agencies (accessed 31 July 2021).