5

The Five Safes
of Risk-Based Anonymization

PRIVACY
ANALYTICS
an IQVIA company

Table of contents

The Five Safes
of Risk-Based Anonymization

Executive summary

Anonymization is a critical piece of the data-sharing puzzle—by its very nature, it enables the responsible sharing of data for secondary purposes.

The sharing of data for the purposes of data analysis and research can have many benefits. At the same time, concerns and controversies about data ownership and data privacy elicit significant debate.

The central question is how to utilize data in a way that protects individual privacy, but still ensures the data are of sufficient quality that the analytics will be useful and meaningful.

When we use the term "anonymization", we mean anonymization that is legally defensible—that is, anonymization that meets and exceeds standards of current legal frameworks, and that can be presented as evidence to governing bodies and regulatory authorities, to mitigate exposure and demonstrate that you have taken your responsibility toward data subjects seriously.

The techniques used to achieve anonymization cannot be separated from the context in which data are shared: the exact data you're working with, the people you're sharing it with, and the goals of subsequent analysis. This is called risk-based anonymization.

There is a framework that has emerged from statistical data sharing by government agencies that is predominantly a risk-based approach. In this white paper, we will demonstrate how it can be operationalized in a broader setting. For the purposes of this discussion, we use "de-identification" as a general term that includes the full spectrum of methods, from simple pseudonymization to full anonymization.

Five Safes

Responsible data sharing requires an assessment of many factors, all of which need to be considered objectively to compare data sharing options. Only then can data custodians determine the most appropriate option for their particular circumstances, given the risks and benefits of sharing data in the first place.

One framework that has gained popularity after more than a decade of use is known as the Five Safes[1], which is intended to capture the relevant dimensions to assess the context and results of a data sharing scenario in an effort to make sound decisions. Those dimensions are: Safe projects, Safe people, Safe settings, Safe data, and Safe outputs.

The advantages of the Five Safes as a model for data sharing is that it can capture a range of options, balancing concerns over the usefulness of the shared data, cost and feasibility of the data sharing scenario, and privacy and confidentiality. The term "safe" is treated on a spectrum, as in "how safe is it?", so that this balancing can take place.

The entire premise of the Five Safes is based on the idea of risk assessment, which may be seen as subjective but with objective support through risk estimation. Greater emphasis is then placed on empirical evidence to drive decision-making.

[1]Desai, T., Ritchie, F., & Welpton, R. (2016). Five Safes: Designing Data Access for Research. *Working Paper. University of the West of England.*

Risk-based anonymization

Risk-based anonymization has been proven to help achieve the balance between protecting individual identity and optimizing data utility. Using this approach, data can be claimed to be anonymous. This is a function of both data transformations, and the additional technical and administrative controls (applied continuously) that are put in place in the process of anonymization. The process is as follows:

- Data transformations are applied once to the data;

- However, the controls need to be applied continuously to ensure the data remains anonymous;

- If the appropriate controls are not in place, lapse, or are not strong enough, then the data is no longer anonymous; the data would then no longer satisfy the scope of privacy or data protection regulation or legislation, and thus be subject to the scrutiny of privacy commissioners or data protection authorities.

Risk-based anonymization requires an evaluation of the external information available to an adversary (whether a re-identification is intentional or not), and how they may combine it to re-identify data. Removing personal information from data using a risk-based methodology requires an assessment of the environment and the circumstances in which the data will be shared (to know what external information will be available to an adversary), and an assessment of the data itself (to determine how the external information available to an adversary may be used to re-identify data).

With that in mind, risk-based anonymization may be stated in terms of the **Five Safes**:

Safe projects:

What are the data flows, is de-identification needed as a privacy-protective measure?

Safe people:

Who are the anticipated data recipients, what are their motivations and capacity to re-identify, and who may they know in the data?

Safe settings:

What are the technical and organizational controls in place to prevent a deliberate attempt to re-identify or to prevent a data breach?

Safe data:

What is the re-identification risk, considering the people and settings of the data environment, and what de-identification can be applied?

Safe outputs:

Would sharing data be perceived as an invasion of privacy to data subjects, are there any ethical concerns with how the shared data will be used?

This document will serve to describe the Five Safes of risk-based anonymization in greater detail.

Safe projects
Know your data flows and understand legal boundaries

Personal data is information about an identifiable individual, often referred to as a data subject. In a data sharing scenario in which we wish to achieve private data analysis, there will always be a sender (the data custodian) and a recipient (the data analyst). But the recipient is also deemed an eavesdropper or adversary (using standard security language). Our goal is to balance the needs of the recipient (providing them with useful data) while minimizing the ability of an adversary (including the recipient) to extract personal information from the data.

It's important to understand the flow of data, to recognize legal boundaries and intended purposes so that we can identify the parameters needed to assess risk and create safe projects.

- Where the collected data are coming from, who collected them, and the legal grounds for doing so.

- Where the shared data are going, who wants access, and the legal grounds for doing so.

- Whether the data are considered personal or not, and how anonymization is applied in accordance to regulations.

Data flow

From source

The data custodian may have collected information for a primary purpose, such as providing care to a patient. Or the data custodian may have collected information explicitly for a secondary purpose, such as constructing a database of patients with diabetes for subsequent research. Personal information may also come indirectly through one or more data custodians, where permitted. Alternatively, data may come from another source claiming to be anonymized (which may need to be assessed in its own right before being used or combined with personal information). Understanding the legal context for collection, approval mechanisms, and transparency, will be important to determine the appropriate mechanisms for sharing data, especially for secondary purposes.

To destination

An agent, acting on behalf of the data custodian, may use personal data for a primary purpose. Depending on the jurisdiction, there may not be a legislative requirement to de-identify information that an agent uses for secondary purposes, or a requirement to obtain additional approval from data subjects for such uses. However, it may be encouraged or desirable. The data custodian may also receive a request to share with an internal or external recipient for some secondary purpose. Sharing of personal data are sometimes mandatory, whereas others may be discretionary to the data custodian. The conditions for discretionary sharing do vary. Other data sharing, that are not explicitly permitted in legislation, require that either consent be obtained from the data subjects or the personal data are anonymized.

When to de-identify

There are four scenarios to consider:

- Mandatory sharing: No approval is required, and the data do not require anonymization because it is likely that individuals need to be identified (e.g., law enforcement). However, there may be considerable underreporting due to privacy concerns.

- Internal sharing: It is often unnecessary for an agent to have data in identifiable form to perform their functions, even for primary purposes, and de-identification is desired to enhance privacy and avoid potential breaches.

- Permitted sharing: Approval may be optional, under the discretion of the data custodian, for the public good (e.g., public health). There is reluctance, however, by data custodians to share personal data due to issues of individual and public trust, which de-identification can help remedy.

- Other sharing: When approval is not possible or practical, and there are no exceptions in the legislation, the custodian must anonymize the personal data before sharing with a data recipient.

2 Safe people
Identify anticipated recipients, and evaluate recipient trust

Data recipients are central to an assessment of context risk because the entity or employees may re-identify data, whether it be intentional or not. The anticipated recipient is also an adversary.

Recipient trust

Consider the motives and capacity of the anticipated data recipient to re-identify the shared data. We assume that the data custodian is sharing data that have gone through some kind of de-identification.

Motives

The motive to re-identify individuals in the data implies an intentional re-identification, considering issues such as conflicts of interest and the potential for financial gain from a re-identification.

Capacity

The capacity to re-identify individuals in the data considers whether the data recipient has the skills and financial resources to re-identify the data.

Motives can be managed by having enforceable contracts with the data recipient. Such an agreement will determine how likely a deliberate re-identification attempt would be. Contractual obligations need to include very specific clauses (otherwise there are some very legitimate ways to re-identify a dataset):

- A prohibition on re-identification, on attempting to contact any of the patients in the data set, and on linking with other data sets without permission from the data custodian;

- An audit requirement that allows the data custodian to conduct spot checks to ensure compliance with the agreement, or a equirment for regular third-party audits;

- A prohibition on sharing the data with other third parties (so that the data custodian can keep track of who has the data), or a requirement to pass on the above restrictions to any other party the data is subsequently shared with.

Acquaintances

Data recipients may have prior knowledge of personal information because they're acquaintances of individuals in the data (e.g., relatives or neighbors). This in turn may lead them to inadvertently, or spontaneously, re-identify data subjects simply by recognizing them. It's a factor that needs to be considered when evaluating risk, because it relates to how safe it is to have people working with data.

The security and privacy practices of the data recipient will have an impact on the likelihood of a rogue employee at the data recipient's site being able to re-identify the shared data. A rogue employee may not necessarily be bound by a contract unless there are strong mitigating controls in place. It also determines the likelihood of an outsider gaining access to the shared data.

An evaluation of mitigating controls needs to be detailed and evidence based, preferably mapped to existing professional, international, and government regulations, standards, and policies, including ISO/IEC 27002, where appropriate. Using a standardized approach also ensures consistency, not only for a single organization that is sharing data, but across organizations, e.g., the HITRUST De-Identification Framework[2].

[2]HITRUST Alliance, (2015). HITRUST De-Identification Framework. Site: hitrustalliance.net/de-identification

Controlling access, disclosure, retention, and disposition of personal data

- Only authorized staff should have access to data, and only when they need it to do their jobs.

- There should be data sharing agreements in place with collaborators and subcontractors, and all of the above should have to sign nondisclosure or confidentiality agreements.

- There should be a data retention policy with limits on long-term use, and regular purging of data to reduce vulnerability to breaches.

- If any data is going to leave the relevant jurisdiction in which the data sharing is taking place, there should be enforceable data sharing agreements and policies in place to control disclosure to third parties.

Safeguarding personal data

- It's important to respond to complaints or incidents, and that all staff receive privacy, confidentiality, and security training.

- Personnel need to be disciplined for violations of these policies and procedures, and there should be a tried and tested protocol for privacy breaches.

- Authentication measures must be in place with logs that can be used to investigate an incident.

- Data can be accessed remotely, but that access must be secure and logged.

- On the technical side, a regularly updated program needs to be in place to prevent malicious or mobile code from being run on servers, workstations and mobile devices, and data should be transmitted securely.

- It's also necessary to have physical security in place to protect access to computers and files, with mandatory photo ID.

Ensuring accountability and transparency in the management of personal data

- There should be someone in a position of seniority who is accountable for the privacy, confidentiality, and security of data, and there needs to be a way to contact that person.

- Internal or external auditing and monitoring mechanisms also need to be in place.

Safe data
Evaluate re-identification risk, incorporating people/settings

An assessment of Safe people and Safe settings results in an evaluation of context risk. A structured approach can be used to assess context risk and evaluate whether an attack will be realized, known as threat modelling.

Consistent with the modelling of threat sources used in information security and risk modelling, there are three plausible attacks that can be made on data:

Deliberate

A targeted attempt by the data recipient as an entity, or a rogue employee due to a lack of sufficient controls, to re-identify individuals in the shared data.

Accidental

An unintentional re-identification, for example an individual being recognized while a recipient is working with the shared data.

Catastrophic

The data could also be lost or stolen in the case where all the controls put in place have failed to prevent a data breach.

To produce Safe data the overall risk of re-identification needs to be assessed, which is a combination of context risk (the probability of an attack) and data risk (the probability of re-identification when there is an attack)[3]. This will drive the de-identification required to reduce risk to an acceptable level.

[3]Skinner, C. J., & Elliot, M. J. (2002). A Measure of Disclosure Risk for Microdata. *Journal of the Royal Statistical Society: series B (statistical methodology), 64(4), 855-867.*

Quantifying risk

Because risk measurement invariably requires the use of statistical methods, any risk measurement technique will be based on a model of plausible re-identification attacks, and models make assumptions about the real world. Therefore, risk measurement will always imply a series of assumptions that need to be made explicit. Furthermore, because of the statistical nature of risk measurement, there will also be uncertainty in these measurements and this uncertainty needs to be taken into account.

Risk metrics

Risk measurement applies to indirectly identifying data. Three kinds of risks need to be managed, of which detailed metrics can be derived[4].

[4]El Emam, K. (2013). Guide to the De-Identification of Personal Health Information. CRC Press.

Prosecutor risk

The prosecutor has background information about a specific person that is known to them, and uses this background information to search for a matching record in the shared data.

Journalist risk

The journalist doesn't know the particular individual in the shared data, which is a subset of a larger public dataset, but does know that all the people in the data exist in a larger public dataset.

Marketer risk

The marketer is less concerned if some of the records are misidentified. Here the risk pertains to everyone in the data. Marketer risk is always less than prosecutor or journalist risk, and istherefore often ignored.

If a population registry has information about individuals who are known to be in the shared data, an adversary may target the highest risk data subjects. In this case, the maximum of the risk metric is taken across all data subjects when there are no controls in place to prevent such an attack (e.g., public data sharing). On the other hand, if an adversary will not target the highest risk data subjects, because there are controls in place to prevent such an attack, but is trying to find information about a specific individual, the risk metric is averaged across all data subjects (e.g., private data sharing).

De-identification

With a measure of overall risk to drive decision making, de-identification must be applied to the data that is intended to be shared. The balance between privacy protection and analytic utility is optimized through the de-identification of indirectly identifying data. Maximum privacy protection (i.e., zero risk) means very little to no information being released. Information loss needs to be minimal so that the data are still useful for data analysis, while ensuring the re-identification risk is very small.

Directly identifying

Attributes that can essentially be used alone to uniquely identify individuals or their households, such as names and known identifiers, are always removed or replaced with pseudonyms. The techniques used need to be robust and defensible.

Indirectly identifying

Attributes that can be used in combination with one another to identify individuals, such as known demographics and events, may need to be modified or transformed to reduce risk (as they are the attributes used to measure risk, and are not eliminated from the shared data).

5 Safe outputs
Consider the potential invasion of privacy, and ethical uses

The degree of de-identification necessary to reduce risk to an acceptable level raises the question of risk thresholds. There are many precedents going back multiple decades for what is an acceptable probability of sharing anonymized data. To decide which threshold to use, we can look at the sensitivity of the data and the approval mechanism that was in place when the data was originally collected.

Invasion of privacy

Invasion of privacy is a subjective criterion that can be used by the data custodian to influence the selection of a risk threshold. If the invasion of privacy is deemed to be high, that should skew the decision more toward a lower threshold. On the other hand, if the invasion of privacy is deemed to be low, a higher threshold would be acceptable.

- Are the data highly detailed, are they highly sensitive and personal in nature?

- What is the potential injury to individuals from an inappropriate processing of the data?

- What is the appropriateness of approval by data subjects for disclosing the data?

Although approval is not required of data subjects for sharing properly anonymized data, the sharing of data would not be considered as privacy invasive when approval has been provided by data subjects compared to when no approval is sought. There are in fact multiple levels of notice and approval that can exist for the sharing of anonymized data.

The practical consequence of evaluating invasion of privacy is that the acceptable threshold (or the definition of "very small risk") will be lower under the most invasive scenario. Even under the most invasive scenario, however, it is possible to share the data, but the degree of de-identification would be greater.

Ethics

Stigmatizing analytics are models that can lead to decisions that adversely affect individuals or groups.

Data custodians who anonymize and share data need to consider the impact of stigmatizing analytics, even though, strictly speaking, it goes beyond anonymization. They should consider building oversight mechanisms and determine how they can meaningfully engender individual trust for the ethical use of data to avoid causing harm, and how they can transfer these obligations to the organizations they share data with.[5]

[5]Richards, N., & Hartzog, W. (2017). Trusting Big Data Research. *DePaul Law Review, 66(2), 10.*

One way to manage risks from stigmatizing analytics is to set up an ethics board that would review analytics protocols. The board would advise and make decisions on whether particular uses of data may be stigmatizing, and whether or how such uses need to be conducted and communicated. This can help the data custodian avoid a spectrum of negative legal, reputational, and regulatory actions.

Internal ethics boards can be beneficial if they are given an appropriate level of authority, and procedures can be put in place to make their degree of involvement proportional to the potential invasion of privacy. It's unlikely, however, that organizations would agree to reveal or uncover confidential commercial information to outside third parties or be bound by external advice.

Conclusion

In many jurisdictions, demonstrating that data has a very small risk of re-identification is a legal or regulatory requirement. Our methodology provides a basis for meeting these requirements in a defensible, evidence-based way. We have demonstrated how the Five Safes framework can be operationalized using risk-based anonymization: each dimension is evaluated independently of the others, brought together by an overall assessment of risk. This allows for the evaluation of scenarios of responsible data sharing, which will be context-driven given the impact different scenarios will have on the usefulness of the data.

Data utility is important for those using anonymized data, because the results of their analyses are critical for informing services provided, policy, and investment decisions. Also, the cost of getting access to data is not trivial, making it important to ensure the quality of the data received. We don't want to be wasteful, spending time and money collecting high-quality data, only to then watch that quality deteriorate through anonymization practices meant to prepare the data for secondary use.

The impact of de-identification on data utility is important, and very context-driven. All stakeholders need to provide input on what is most important to them, be it data utility or privacy. It's not easy to balance the needs of everyone involved, but open communication and a commitment to producing useful data with a sufficiently low risk of re-identification is all that is really needed to get started. It's not an easy negotiation—and it may be iterative—but its importance cannot be underestimated. Ideally, framing that conversation around the Five Safes should help to clarify the most important points.

Authors

Khaled El Emam
CEO

kelemam@privacy-analytics.com
linkedin.com/in/kelemam

Luk Arbuckle
Chief Methodologist

larbuckle@privacy-analytics.com
linkedin.com/in/lukarbuckle

Related reading

El Emam, K., & Arbuckle, L. (2014). *Anonymizing Health Data: Case Studies and Methods to Get You Started. O'Reilly Media, Inc.*.

Stalla-Bourdillon, S., & Knight, A. (2016). Anonymous Data v. Personal Data-False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data. *Wis. Int'l LJ, 34, 284.*

Polonetsky, J., Tene, O., & Finch, K. (2016). Shades of Gray: Seeing the Full Spectrum of Practical Data De-Identification. Santa Clara *L. Rev., 56, 593.*

Appendix: Summary of principles

1

Safe projects
- Know your data flows and understand legal boundaries

2

Safe people
- Identify anticipated recipients, and evaluate recipient trust

3

Safe settings
- Assess the security and privacy practices of the recipient

4

Safe data
- Evaluate re-identification risk, incorporating people/settings

5

Safe outputs
- Consider the potential invasion of privacy, and ethical uses

About Privacy Analytics

Our methodology, based on best practice, was developed out of the Electronic Health Information Laboratory (EHIL) headed by Khaled El Emam, PhD. Formalized in 2005 under the Children's Hospital of Eastern Ontario (CHEO) Research Institute in Ottawa, Canada, EHIL is the only research group in Canada conducting both theoretical and applied research on the anonymization of health information and secure computation over health data, and has produced many qualified researchers.

Incorporated in 2007, Privacy Analytics, headquartered in Ottawa, is a spin-off from EHIL. Privacy Analytics' mission statement includes the enablement of healthcare and non-healthcare organizations to responsibly apply a risk-based anonymization methodology, in order to safeguard individual privacy, while meeting and exceeding standards for legal compliance. Privacy Analytics, following its acquisition by IMS Health (now IQVIA), is an autonomous subsidiary of IQVIA, a leader in Real World Evidence solutions for Life Sciences, with a network in 100 countries.

As of the writing of this whitepaper, Privacy Analytics occupies a unique position in the data-privacy industry as a single-source provider, offering expert training, software, peer-reviewed methodology and valued-added services; to protect the privacy of individuals' personal and health care information, while enabling organizations to share this data for secondary purposes.

Privacy Analytics' client base includes more than half of the healthcare companies in the Fortune 50. Our team of data-privacy experts are known as pioneers in the development of methodologies, software and services that enable responsible use of complex data assets comprising personal information. Privacy Analytics experts are frequently solicited to present at speaking engagements, industry events and in academic circles around the world. Our advice is also sought by several entities who shape data-privacy policy.

To learn more, please visit: www.privacy-analytics.com

Privacy Analytics Inc.

251 Laurier Avenue West, Suite 200
Ottawa, Ontario, Canada K1P 5J6

Telephone: +1 613 369 4313

Toll free: +1 855 686 4781

Fax: +1 613 369 4312

Email: marketing@privacy-analytics.com

Visit us on the web: www.privacy-analytics.com