



**Office of the Victorian
Information Commissioner**

A blurred, high-angle photograph of a crowd of people walking on a light-colored tiled floor, overlaid with a semi-transparent grey filter.

De-identification and the IPPs

An exercise in risk management

Emma Stephens
Senior Policy Officer

Bryan Wee
General Counsel

Deidentification

An existential question, a warning and a parental advisory.

Section 3 of the PDP Act:

De-identified, in relation to personal information means personal information that no longer relates to an identifiable individual or an individual who can be reasonably identified

The existential question

Do you need to “use” information under IPP2 in order to “deidentify” information under IPP4?

A warning – can you do it at all?

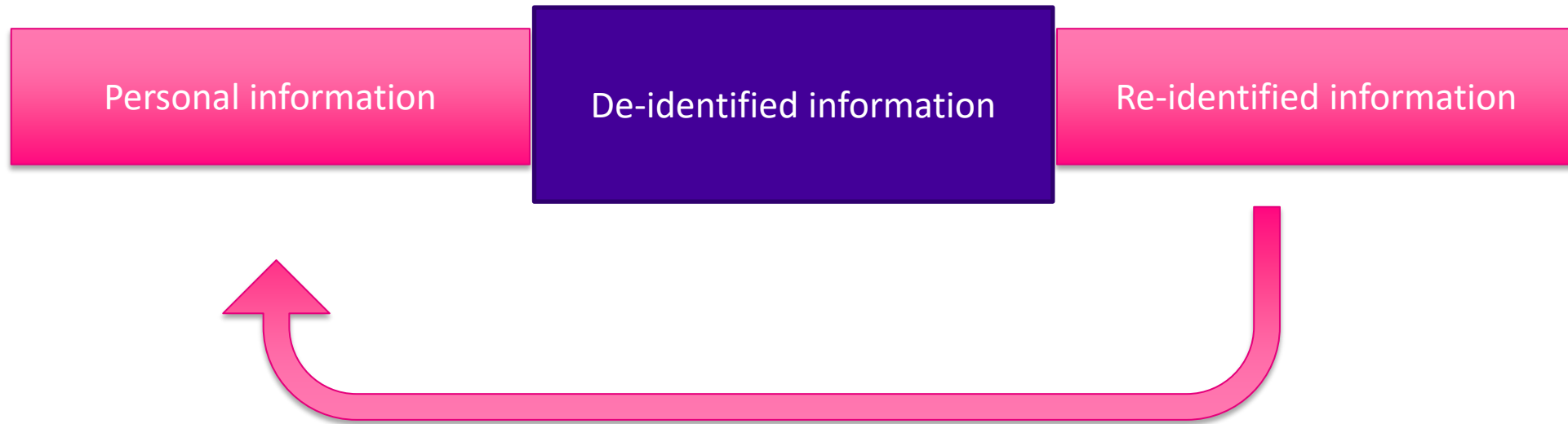
- Think about your enabling legislation:
 - use restrictions
 - secrecy/confidentiality provisions
- Contractual obligations

Parental advisory

- When you adequately deidentify information, the PDP Act and the IPPs no longer apply to it: see sections 16, 19 and 20 for example.
- (This may not be the case for your enabling legislation).
- However, when you release deidentified information, you remain responsible to ensure it remains deidentified and cannot become personal information again because of release of new information (a Rosetta stone), new datasets or new technology.
- Much like having children, deidentification is filled with worry, angst, responsibility but can be very rewarding.

How to assess if information is de-identified

Re-identification

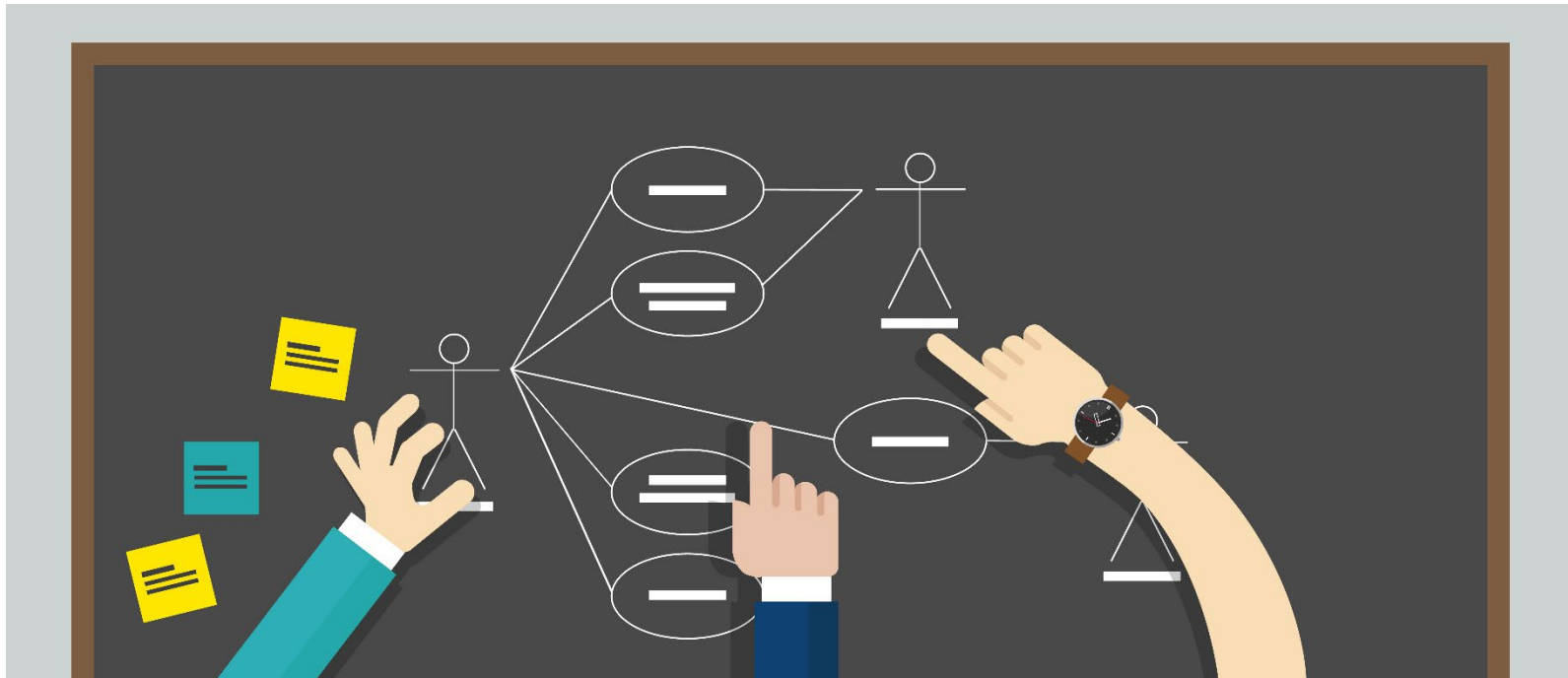


When is data de-identified?

When the risk of an individual being **re-identified** in the **data access environment** is **very low**.

It's about more than a de-identification technique

Context is key



© bakhtiarzein / Adobe Stock

Risks are ongoing...

Consider how combinations of data could be used to single a person out

- Salary
- Superannuation

- Position title
- Years spent working in the position

Name

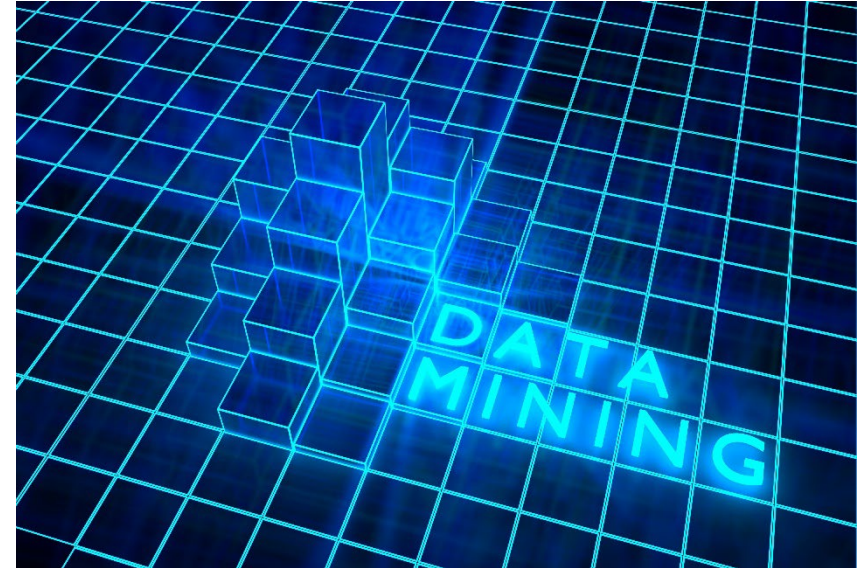


Consider who will have access and how the data will be used

Risk of spontaneous recognition



© metelsky25 / Adobe Stock

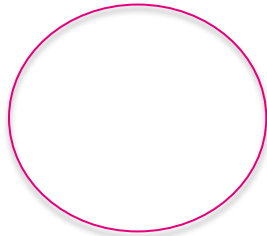


© beebright / Adobe Stock

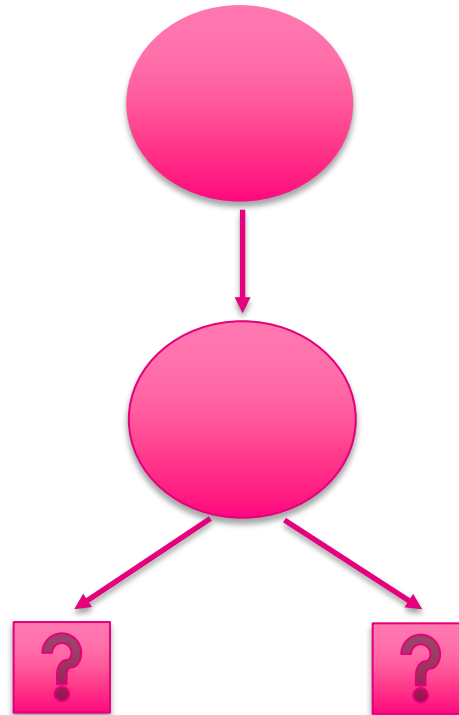
Inferring personal or sensitive attributes

Consider the data access environment

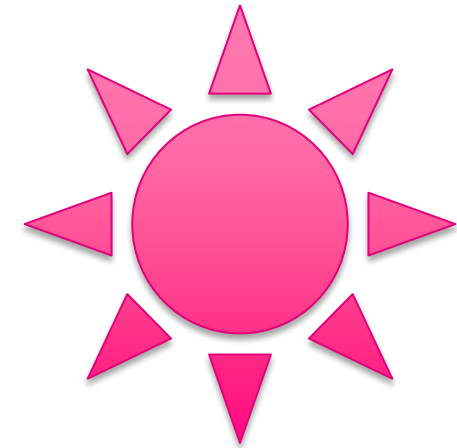
Closed

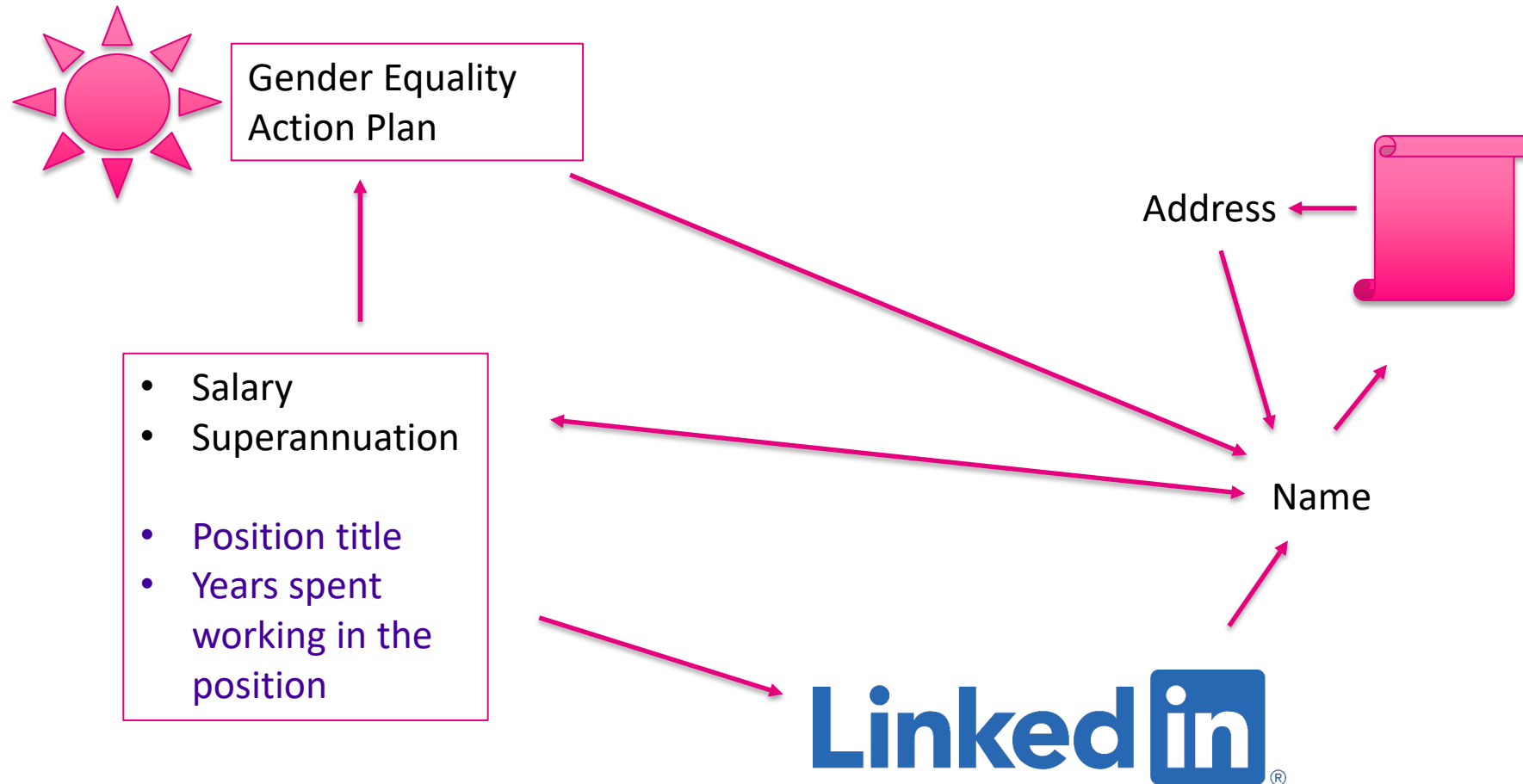


Third party sharing



Open to the public





Consider whether a motivated adversary could identify an individual

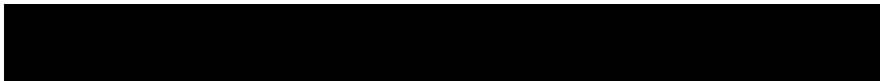
The motivated adversary is a person:

- who is reasonably competent and motivated
- with commonly available skills
- with access to the information, the internet and public documents
- who would make reasonable enquiries to gain more information.



If the motivated adversary could single a person out, the information is not de-identified.

Consider the techniques or methods used to de-identify the personal information

- What technique is appropriate for the situation?
- What will reduce the risk of re-identification to very low and ensure the information remains useful?
- Can synthetic data be used instead?
- Redaction 
- Replacing narrow words with general words

Consider the techniques or methods used to de-identify the personal information

- What technique is appropriate for the situation?
- What will reduce the risk of re-identification to very low and ensure the information remains useful?
- Can synthetic data be used instead?
- Redaction **I'm so good at redacting! The best, really.**
- Replacing narrow words with general words

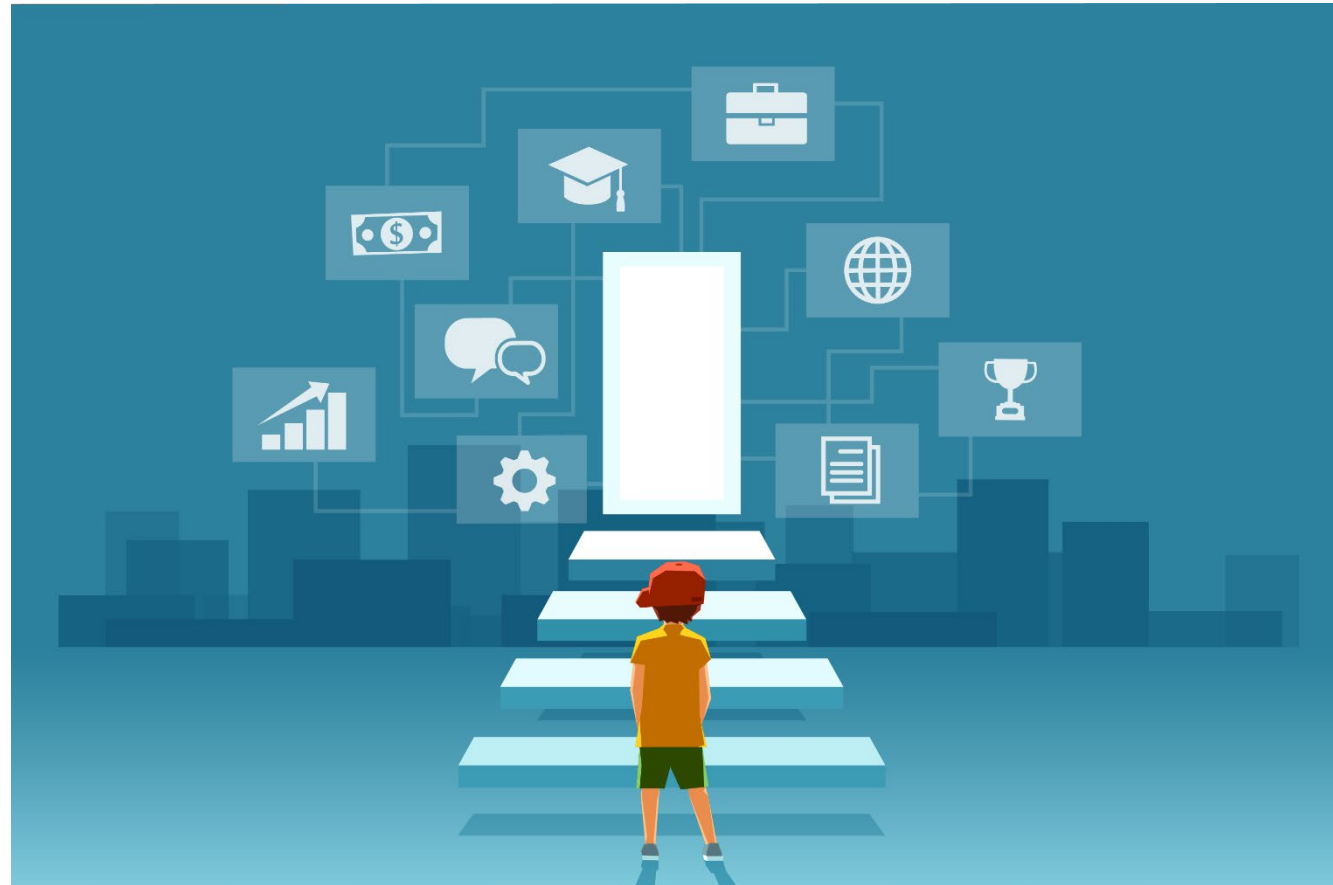
De-identification techniques

- K-anonymity / frequency rule
- Differential privacy
- Removing variables unique to the individual
- Encryption

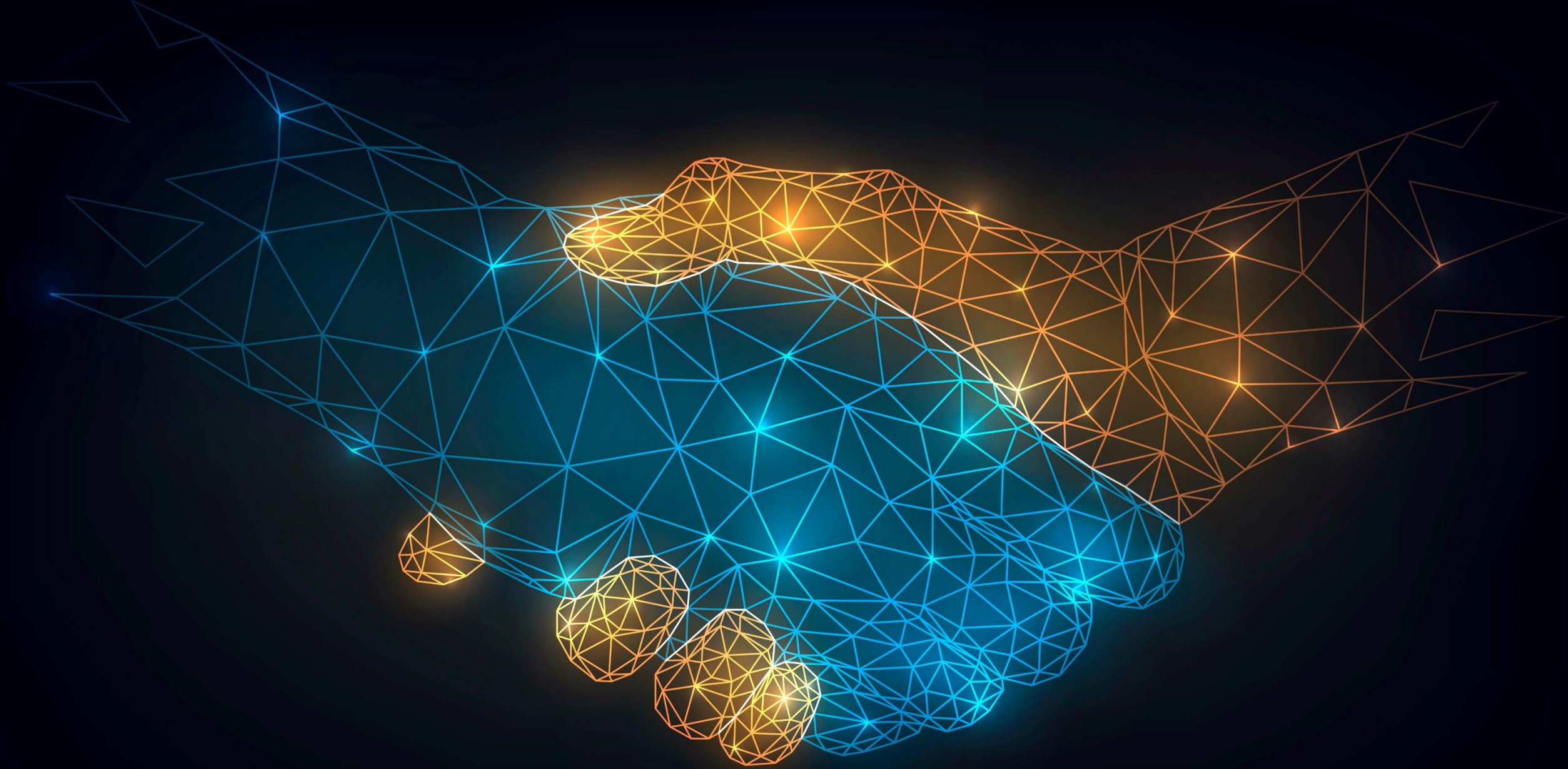


© AliFuat / Adobe Stock

Consider the risk of harm if a person was identified

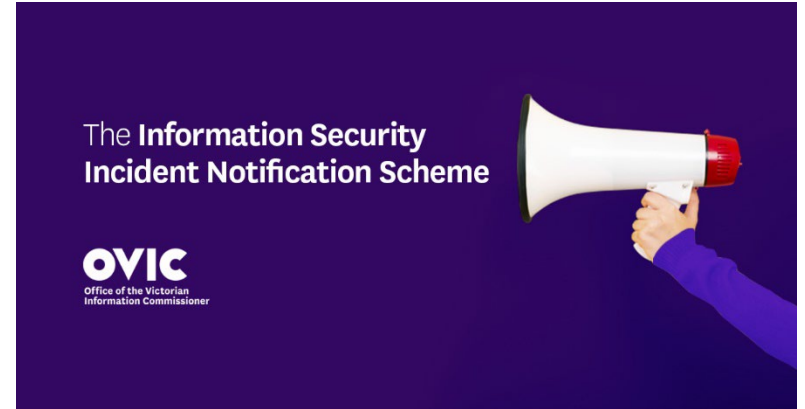


© Feodora / Adobe Stock



What do I do if information becomes re-identified?

1. Remove third party access to the information.
2. Consider whether it is a **reportable incident** and follow your organisation's **data breach response plan**.
3. Consider how the risk of re-identification could be reduced to **very low** to allow the information to be shared.
4. If the risk cannot be reduced to very low, **do not disclose** the information.
5. Use it as a learning opportunity - **what might you do differently next time?**



Further information

- OVIC : [An introduction to de-identification](#)
- OVIC: [Limits of de-identification](#)
- OAIC and CSIRO Data 61: [De-identification and Decision-Making Framework](#)
- [Five Safes Framework](#)

