# Privacy for Infosec Pros

## AusCERT 2015 Tutorial
## 2 June 2015

**Stephen Wilson**
**Lockstep Group**

# Master Class: Objectives & Agenda

*Bring out some privacy surprises*

*Explain data privacy mechanics*

*Surface some of the tensions*

*Help resolve them*

*Re-position infosec and privacy.*

1. Recap: Data Privacy Regulations
2. Bridging Infosec and Privacy
3. Case studies – interspersed
4. Practical *Privacy Engineering*
5. Summary and Discussion.

# Privacy in plain language

Lockstep suggests that Privacy is a state where a party that has Personal Information about you is restrained in how they use that information.

**Privacy ≠ Secrecy**

# Australian Privacy Principles – highlights

## Personal Information

> information or an opinion about an identified individual, or an individual who is reasonably identifiable ... whether the information or opinion is true or not ...
>
> *Privacy Amendment (Enhancing Privacy Protection) Act 2012*

## APP 3 Collection

An organisation must not collect personal information unless the information is reasonably necessary for one or more of the organisation's functions or activities.

## APP 6 Use and disclosure

Personal information collected for a particular purpose must not be used or disclosed for another purpose unless the individual consents.

## APP 1 [Openness]

Must set out clear policies about how personal information is managed; generally what sort of personal information is held, why, and how it is handled.

## APP 3.3 [Sensitive Personal Information]

Personal Information about health, race, ethnicity, religion, sexuality, politics etc. must not be collected without informed consent.

# Privacy is a Technology Issue

| | |
|---|---|
| 🟨 APP 1: | Open and transparent management of PI |
| 🟨 APP 2: | Anonymity and pseudonymity |
| 🟥 APP 3: | Collection of solicited PI |
| 🟥 APP 4: | Dealing with unsolicited PI |
| 🟨 APP 5: | Notification of the collection of PI |
| 🟥 APP 6: | Use or disclosure of PI |
| APP 7: | Direct marketing |
| 🟨 APP 8: | Cross-border disclosure of PI |
| 🟨 APP 9: | Adoption, use or disclosure of govt identifiers |
| 🟥 APP 10: | Quality of PI |
| 🟥 APP 11: | Security of PI |
| 🟥 APP 12: | Access to PI |
| 🟥 APP 13: | Correction of PI. |

It's often said, in a well intended way, that "privacy is not a technology issue", meaning that all sorts of things go into privacy solutions. Technologists must not think it has nothing to do with them, for almost every one of the Australian Privacy Principles may be impacted by technology to a moderate (yellow) or a major extent (red).

# Privacy Management

- ## Privacy Policy

  A Privacy Policy is typically written by lawyers. Unfortunately, because of the compliance focus, and in some cases, due to business practices that seek to exploit personal information, Privacy Policies have become more like legal disclaimers. Lockstep advises that a genuine Privacy Policy should focus on restraint and openness. It should set out plainly What PI is collected, Why it is collected, How, Where and When it is collected, and Who else will see the PI. In addition to explaining what is done with PI, an organisation should explain what it will not do.

- ## Privacy Impact Assessment (PIA)

  A Privacy Impact Assessment is a semi-standardised way of understanding what a program or information system will do to privacy, by mapping the major personal information flows and analysing them against the relevant Privacy Principles. More later.

- ## “Privacy By Design” (PbD)

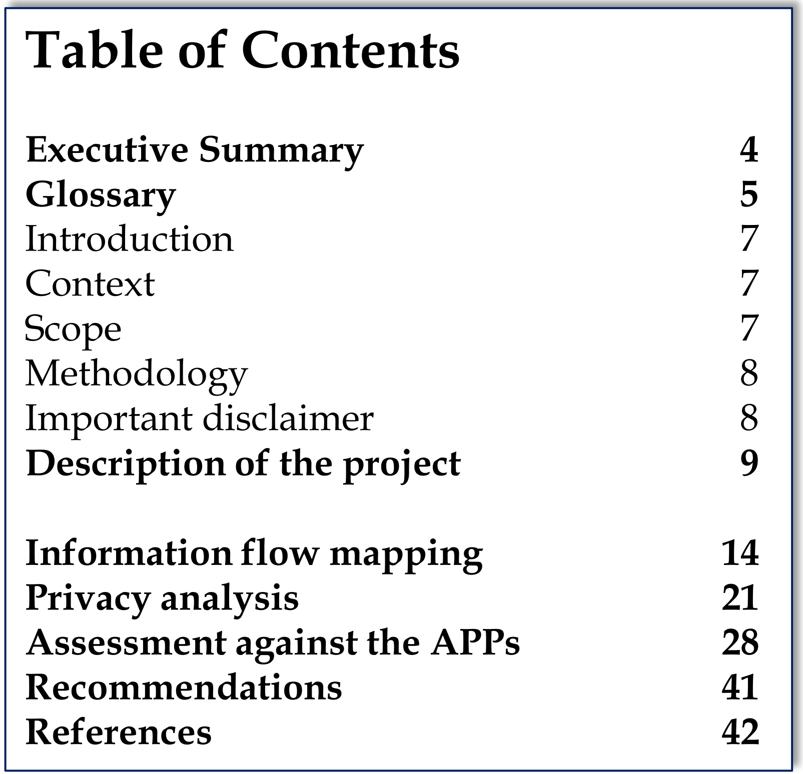
  *Privacy by Design* is a movement that promotes early consideration of privacy by projects, and integration of positive privacy attitudes into design and development. PbD means well but it is seldom more than a simple re-jigging of technology neutral privacy principles, with some frankly unhelpful idealistic overlays. More later.

- ## Privacy Engineering

  Privacy Engineering is a newer idea, being explored by various researchers, to orientate IT practitioners to privacy. Lockstep’s work in this area aims to expose and resolve the inherent tensions in privacy, much as we do with security-cost or security-performance tradeoffs. More later.

# Privacy Impact Assessment (PIA)

- The PIA is semi-standardised in Australian state and federal government, and is increasingly practiced by business.
- It is important to distinguish PIAs done for checking compliance with regulation s (an important exercise, towards the launch of a service) from ones done at the formative stages of the project, as a design tool. Both are important.
- We feel the best PIA guidelines are probably those of Privacy Victoria and OAIC.
- A typical table of contents is shown.



## Table of Contents

# Technology Neutrality

- “Collection” is not defined in most data privacy legislation. A technology neutral view is that whenever and however personal information ends up in a system, as a matter of logic, it has been *collected*, and privacy responsibilities arise.
- The words “public”, “private” and “public domain” are not operable in the Privacy Act! Our law is largely blind to where PI comes from.
- “Collect” is not a *directive* verb. You don’t need to *actively* collect personal information; if it comes to you, you have collected it.
- Therefore it can be useful to think in terms of *Synthetic PI*, like the outputs of Big Data when identifiable, and *Algorithmic Collection,* as opposed to direct collection.

# Case: Street View Wi-Fi

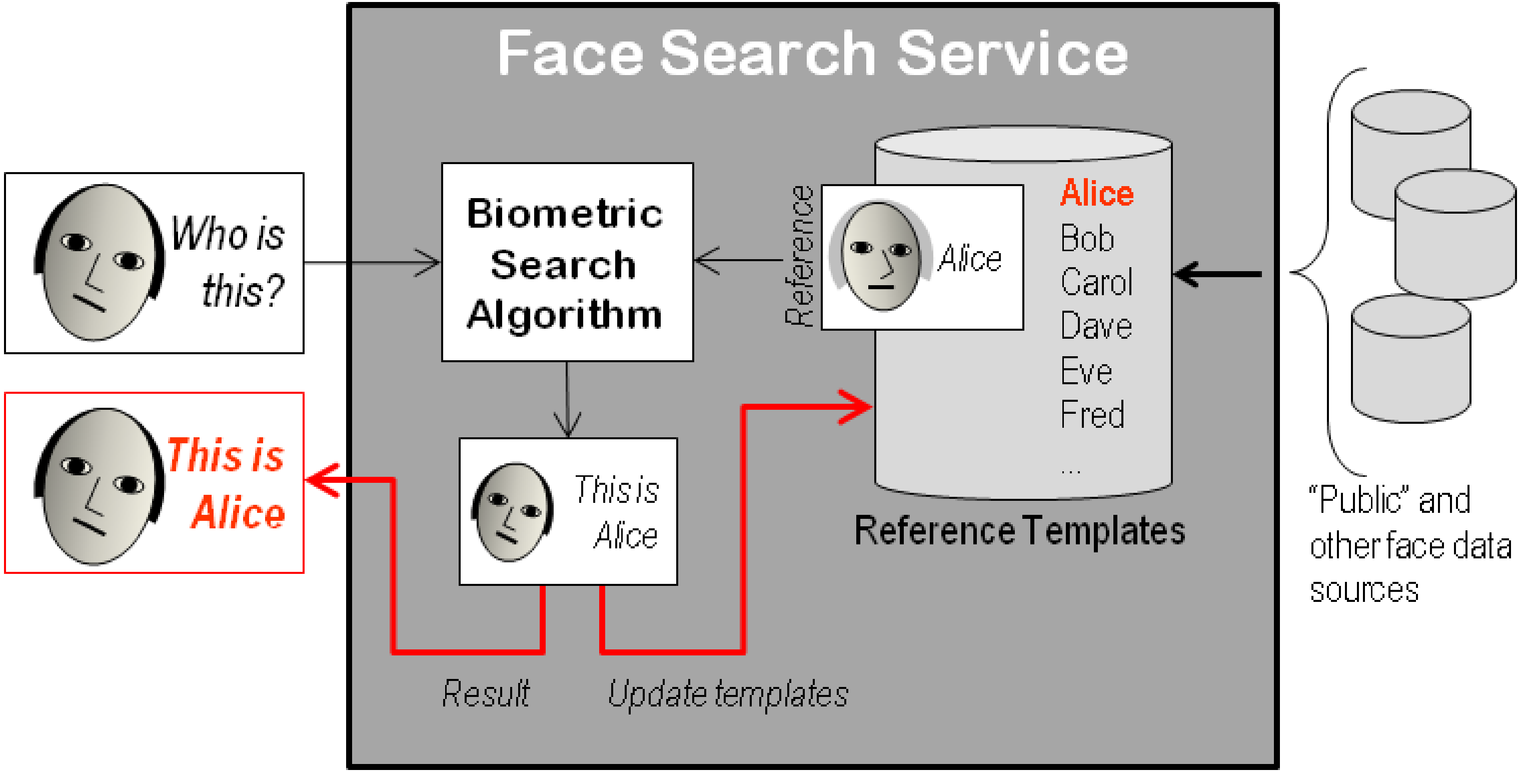
- SSIDs are collected by Google Street View cars to build up a global geo-location database. Privacy impact is negligible.
- But the cars “accidentally” collected Wi-Fi contents too from unencrypted  wireless networks.
- Unencrypted data may be identifiable.
- That the network data was in the “public domain” is irrelevant in most countries with technology neutral data privacy laws.
- Australia, European and Asian countries responded swiftly.
- The US response is bogged down in complex telecomms law.

# Case: Facial Recognition

- Biometric templates are generated when a photo is tagged.
- Facial recognition creates *tag suggestions*.
- Converting anonymous photos into identified ones is a synthetic collection of personal information .
- European regulators shut down tag suggestions, as it was a form of collection without permission.
- In Australia, biometrics for identification is now classed as *Sensitive* PI and must not be collected without informed consent.

# Case: Face Search

Face search services (like the controversial *Name Tag App*) might use reference material drawn from the public domain, but the fact remains they create fresh PII (in red) and as such are accountable under standard data privacy laws.



# Common ground between security and privacy

- Collection Limitation
- Least Privilege
- Need to Know

*Must not collect PI unless reasonably necessary for one or more functions or activities.*

*Entities (people, processes, devices) should be assigned the fewest privileges consistent with their assigned duties and functions.*

# Privacy by Design (PbD) Foundational Principles – a critique

1. ***Proactive not Reactive; Preventative not Remedial***
2. ***Privacy as the Default Setting***
3. ***Privacy Embedded into Design***
4. ***Full Functionality – Positive-Sum, not Zero-Sum***
5. ***End-to-End Security – Full Lifecycle Protection***
6. ***Visibility and Transparency – Keep it Open***
7. ***Respect for User Privacy – Keep it User-Centric***

The first three principles border on motherhood. The same sorts of things are said about building in security, or building in quality. On their own, slogans mean little, and often give way to other imperatives.

The “Positive-Sum” (No. 4) is a noble idea; we can and should have privacy and security together. Yet there *are* inherent tensions. It is naive to simply paper over the notorious “Zero Sum” anti-privacy slogan.*

No. 5 is a rewording of standard privacy (eg APP 11).

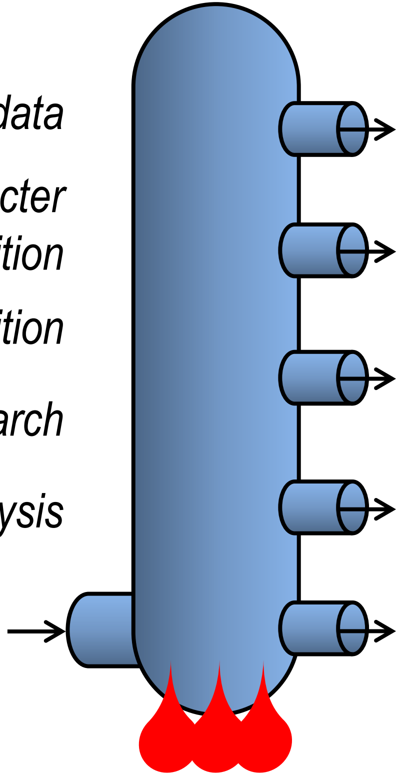
No. 6 is a rewording of standard privacy (eg APPs 1, 12 & 13).

User centricity in privacy (No. 7) is usually taken to mean user *control*, but this is especially problematic in the era of Big Data and IoT, where the vast majority of PI is collected or synthesised behind our backs, beyond our control.

** "We have a saying in this business: 'Privacy and security are a zero-sum game'" - Ed Giorgio, NSA, 2008.*

# Big Data as Crude Oil

Data is often compared to crude oil. It's a powerful metaphor, especially when we consider the long and complex supply chains involved. We refer to "data mining" but data *refining* is really where the action is. The digital economy is bringing about all sorts of new data based value chains. When personal information is extracted from raw data, privacy obligations are inescapable



*Picture metadata*

*Character recognition*

*Face recognition*

*Object search*

*Facial analysis*

**Photo data**

**Location**
*from placenames, landmarks*

**Linkages**
*from who took photo, recognised companions*

**Deduced Likes**
*from trend data, recognised objects, emotions*

**Behaviour Patterns**
*from trend data*

**Future Intelligence?**

# Case: Pregnancy predictor

- A Big Data classic! Target researched and developed algorithms to mine women’s shopping habits to find out if they are pregnant.
- This sort of Big Data Avoids direct collection of personal data, and yet consumers expect businesses to be restrained in how they handle intimate personal data – especially when it’s collected covertly.
- In Australia, Health Information like the state of pregnancy is *Sensitive* and must not be collected (including synthesised) without express consent.
- Thus Big Data challenges the *Collection Principle*: It is tricky to set out in a privacy policy today the sorts of personal data that might be synthesised tomorrow.

# Case: The “Right to be Forgotten”

- A 2014 EU court decided that people affected by wrong or out dated web search results have the right to have results delisted by the search provider.
- “RTBF” is a misnomer. De-listing doesn’t affect original material; there is no forgetting, nor any affect on curtail original sources.
- Some say it’s censorship or “rewriting history”, but what is “truth”? Search is a Big Data process. The results are a highly contextualised prediction of what the user is really interested in. Search results vary from place to place and day to day.
- A classic case of individual vs corporate rights. RTBF provides for search engines to be more responsive to end user needs.

# Trade data not privacy!

- “Trade privacy” is a category error
- Privacy is not a thing; *data* is a thing
- Trade services for data, preserving privacy
- Trade goods for money, preserving value
- Social Media PI bargain is fine in principle
- Openness, fairness, negotiation.

# Privacy Engineering

Engineering is all about resolving design compromises in a systematic and agreed manner. Real world tensions are everywhere, for example: Cost vs Performance , Performance vs Functionality, Functionality vs Security, Security vs Convenience , and Security vs Cost.

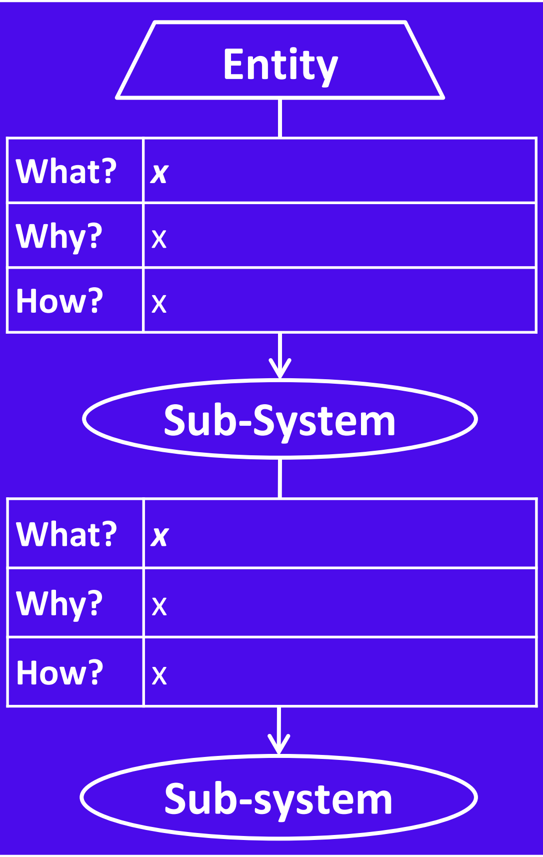
If we are serious about "Privacy engineering" then we need to deal with:

*Privacy vs Convenience*

*Privacy vs Security*

*Privacy vs New revenue*

# Personal Info Flow Mapping



Entity

| | |
|---|---|
| What? | *x* |
| Why? | x |
| How? | x |

Sub-System

| | |
|---|---|
| What? | *x* |
| Why? | x |
| How? | x |

Sub-system

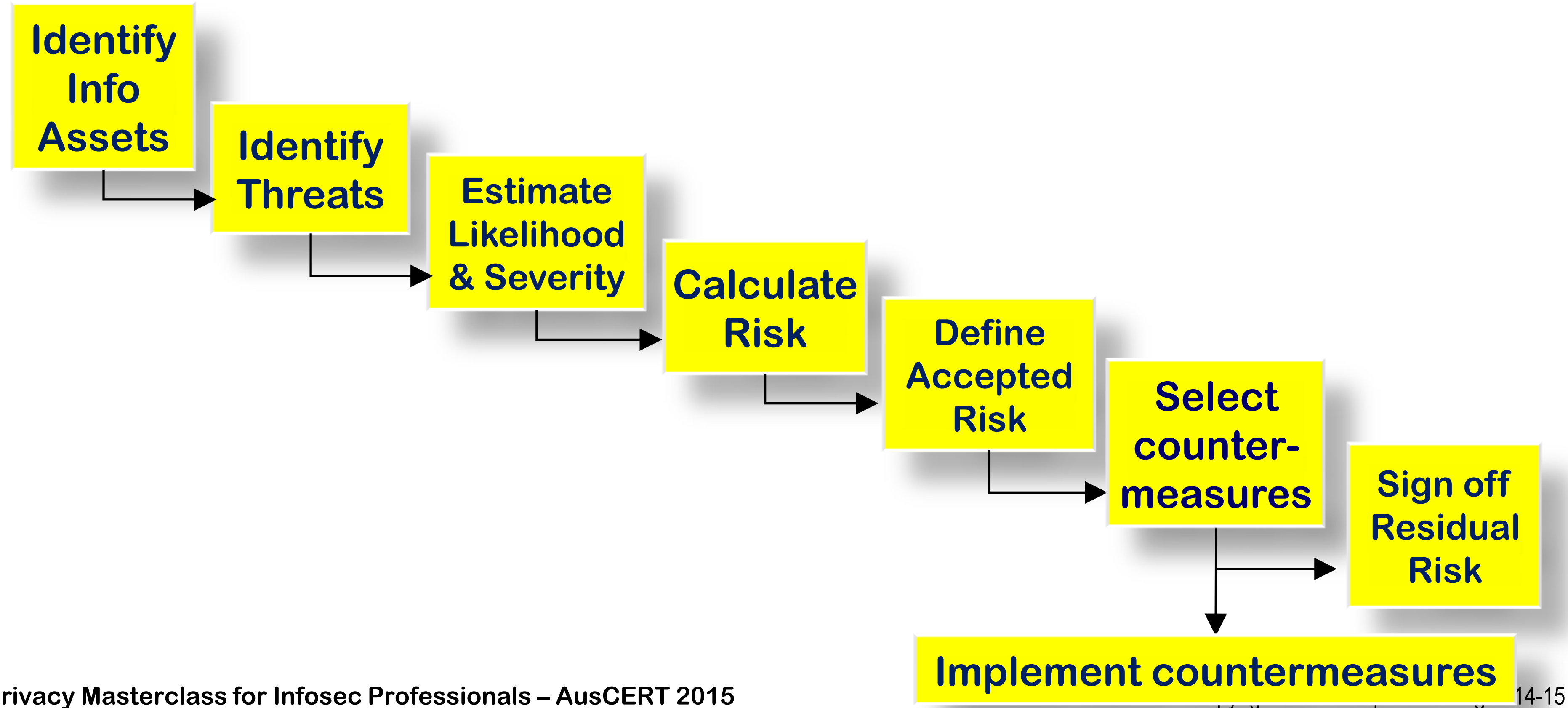
Lockstep's Personal Information Flow Maps label all major data flows with their privacy-relevant qualities. For each data item, we describe Why it is collected or disclosed, How collection / disclosure happens, Where the data comes from and Where it goes.

# Extended Information Inventory

| Item | C | I | A | Privacy factors | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Personal? | Transborder flow? | What permissions? | Sensitive? | Source of data? |
| *Customer record* | | | | | | | | |
| *Employee record* | | | | | | | | |
| *Price list – products* | | | | | | | | |
| *Price l* | | | | | | | | |
| *Acces* | | | | | | | | |
| *Acces* | | | | | | | | |
| *Commerce server event log* | | | | | | | | |
| *Firewall rule set* | | | | | | | | |

A conventional Information Asset Inventory lists the Confidentiality, Integrity and Accessibility of critical data items. The extended Privacy Asset Inventory notes whether each item is *Personal Information*, and if it is, whether the data may flow across borders, if it has been consented for e.g. Direct Marketing, and records the Source(s) of the Personal Information.

# Classic Threat & Risk Assessment



Identify Info Assets

Identify Threats

Estimate Likelihood & Severity

Calculate Risk

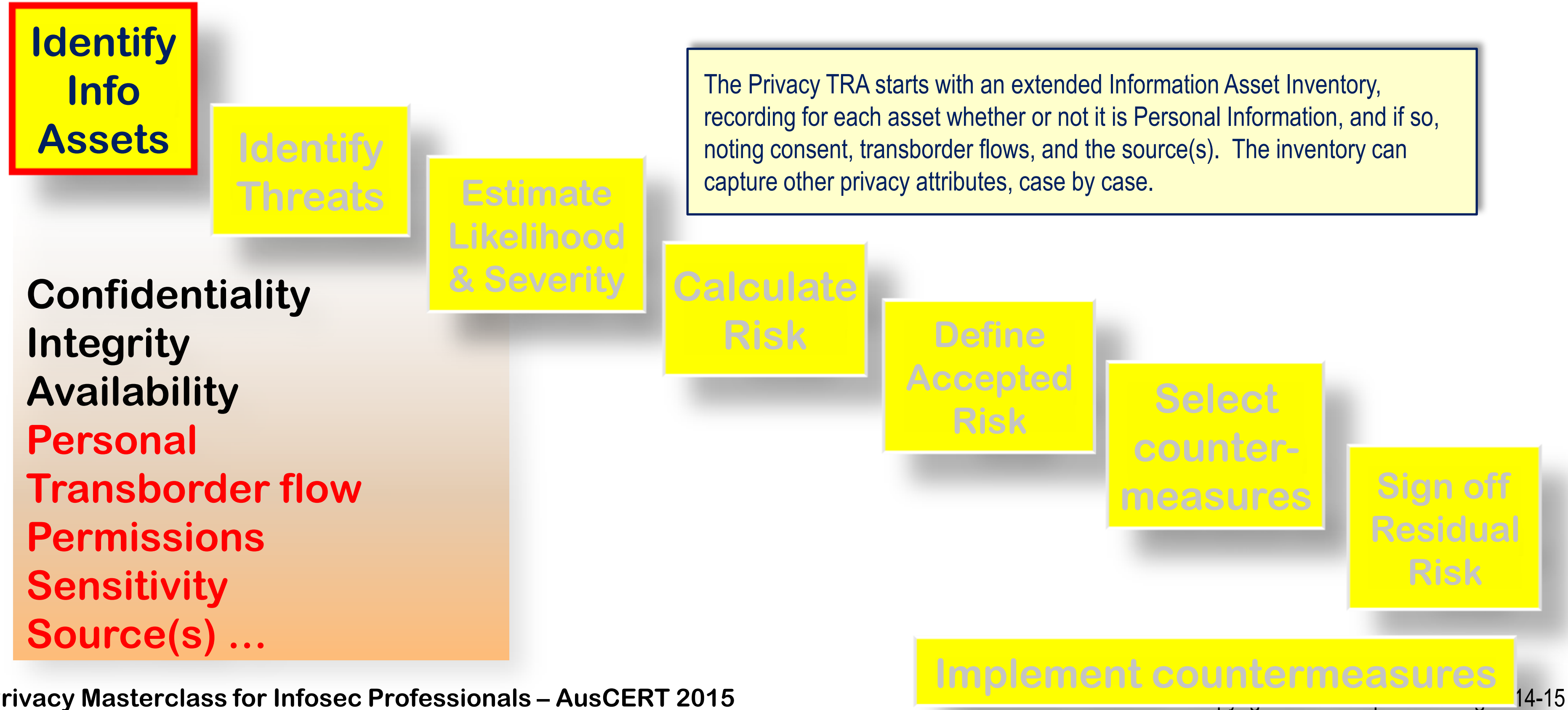
Define Accepted Risk

Select counter-measures

Sign off Residual Risk

Implement countermeasures

# Hybrid Threat & Risk Assessment



Identify Info Assets

Identify Threats

Estimate Likelihood & Severity

Calculate Risk

Define Accepted Risk

Select counter-measures

Sign off Residual Risk

Implement countermeasures

Confidentiality
Integrity
Availability
Personal
Transborder flow
Permissions
Sensitivity
Source(s) …

The Privacy TRA starts with an extended Information Asset Inventory, recording for each asset whether or not it is Personal Information, and if so, noting consent, transborder flows, and the source(s). The inventory can capture other privacy attributes, case by case.

# Hybrid Threat & Risk Assessment

Identify Info Assets

**Identify Threats**

Destruction
External hack, inside job etc.
Over Collection
Unwarranted Use
Wrongful Disclosure
Over retention
Re-identification …

Estimate Likelihood & Severity

Calculate Risk

Define Accepted Risk

Select counter-measures

Sign off Residual Risk

Implement countermeasures

A classic Security TRA draws on a common catalogue of threats such as destruction of data, hacking etc. The Privacy TRA similarly operates with a set of common privacy threats, such as unwarranted collection of PI, unwarranted retention, the compliance of partners to which PI is disclosed, and, very topically, the potential for "anonymised" data to be re-identified. .

# Hybrid Threat & Risk Assessment

Identify Info Assets

Identify Threats

Estimate Likelihood & Severity

Calculate Risk

Define Accepted Risk

**Select counter-measures**

Sign off Residual Risk

Access control *
Encryption *
Personnel *
Event logging *
User Interface
Metadata
De-identification …

The final step of the hybrid Privacy TRA is to select countermeasures and re-calculate the residual risks. Many privacy risks may be mitigated using standard security countermeasures, indicated by red asterisks. Additional privacy-specific measures can include new user interfaces for capturing informed consent, metadata for managing the source and permissions for each data item, and contract terms that commit partners to the APPs as applicable.

**Implement countermeasures**

# Food for thought

- Technology neutrality
- Direct Collection and Indirect Collection
- Privacy is an information asset
- There is no perfect privacy
- Nothing is certain.

# Take aways

- Privacy per se doesn't stop you from handling information
- Privacy is less about what you do than what you don’t do
- If Big Data purposefully avoids direct collection
  why should privacy responsibilities be any different?
- Digital photos should probably be regarded as Personal Info
- Data *ownership* is irrelevant in Australian privacy law
- We can and should trade Personal Information for services, if the bargain is fair and open, then we don’t lose privacy as a right
- Most “privacy enhancing” technologies are secrecy tools
- Privacy is the protection you need when you’re not anonymous.

# Privacy is …

- political
- not much about security
- not at all about secrecy
- all about restraint
- not to be sugar coated.

# Resources

- *'I've Got Nothing to Hide' and Other Misunderstandings of Privacy*, Daniel Solove 2007 http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=998565
- *A Theory of Creepy: Technology, Privacy and Shifting Social Norms*, Omer Tene & Jules Polonetsky 2013 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2326830
- Prof. Graham Greenleaf http://www2.austlii.edu.au/~graham/
- *Global Tables of Data Privacy Laws and Bills*, Greenleaf 2015 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2603502
- *Privacy Impact Assessments guide* – Privacy Victoria 2009 https://www.privacy.vic.gov.au/privacy/web2.nsf/files/privacy-impact-assessments-guide

# Resources

- Lockstep Privacy Library http://lockstep.com/au/library/privacy
- Lockstep Privacy Blog http://lockstep.com/au/blog/privacy
- *Privacy Enters Adolescence*, Constellation Research, 20 December 2014 at https://www.constellationr.com/content/state-state-privacy-enters-adolescence
- *The Apple Privacy Watch* Constellation Research, Sep 2014 https://www.constellationr.com/content/apple-privacy-watch
- *It's not too late for privacy,* Constellation Research, Oct 2012 https://www.constellationr.com/content/its-not-too-late-privacy
- *Siri: A penny for your thoughts?* Lockstep, Mar 2012 http://lockstep.com.au/blog/2012/03/12/a-penny-for-your-thoughts
- *What stops Target telling you're pregnant?* Lockstep, Mar 2012 http://lockstep.com.au/blog/2012/03/07/target-tells-youre-pregnant