REVEALING PRIVACY REQUIREMENTS

HONOURING THE SOCIAL IMPERATIVE IN THE BA'S WORK

FOUNDATIONAL PRINCIPLES OF PRIVACY

- Principle 1: Proactive, Not Reactive; Preventative, Not Remedial
- Principle 2: Privacy as the Default Setting
- Principle 3: Privacy Embedded into Design
- Principle 4: Full Functionality — Positive Sum, Not Zero Sum
- Principle 5: End-to-End Security — Full Life Cycle Protection
- Principle 6: Visibility and Transparency — Keep It Open
- Principle 7: Respect for User Privacy; Keep It User Centric

7foundationalprinciples.pdf (ipc.on.ca)

CANADIAN PRIVACY LAWS

- **BILL C-27,** tabled by the Gov't of Canada on June 16, 2022, titled: *"An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts.*
- The Bill is designed to update Canada's federal private sector privacy law, the *Personal Information Protection and Electronic Documents Act* (**PIPEDA**), to create a new tribunal, and to propose new rules for artificial intelligence (AI) systems.

Source: Dentons - Canada's new federal privacy Bill C-27 – Summary of significant impacts and new proposals

FORCES ENGENDERING THE NEED FOR PRIVACY PRINCIPLES

1. Structural Asymmetry
   - Asymmetric information & power
   - The “bandwidth problem”
2. Cognitive Biases
   - The “endowment effect”
   - Hyperbolic discounting [also expr. In section 5.2.2.2, pg. 179/419]

DOMAIN / THREAT ACTORS

- Any person who interacts with an individual or their information represents a potential threat.
- This is true even for those who have legitimate “need to know” – they could overstep their mandate (e.g. corruption, curiosity)
- Which is why we need to ensure proper roles and permissions are in place; hence the overlap between information security & privacy.
- **Capability vs. capacity:** the former refers to the skills and resources at a threat actor’s disposal, whereas the latter has to do with what means the actor has at their disposal *given their position in the system*.
  - So, a seasoned hacker would be high capability, but a file clerk with improperly elevated permissions would be high capacity. Either way, an astute business analyst would do well to group functions requirements based on classes of legitimate actors, and exclude illegitimate actors from such a matrix.

POTENTIAL FALLOUT OF PRIVACY "HOLES"

		Capability	
Motive	**Scenario**	**Less Capable**	**More Capable**
Commit a crime	Accessing your email account and then using information found to blackmail you. (**Blackmail**)	*Lone hacker*	*Organized criminals* More resources to hack accounts.
Make money	Aggregating data about you to try to sell you more services. (**Aggregation**)	*Shopkeeper*	*Multi-national* Economies of scale allow the multi-national company to get more data.
Enforce the law	Using a fingerprint found at a crime scene to identify you.[5] (**Identification**)	*Local police*	*FBI* More resources at its disposal.
Inflict harm	Posting nude photos of you on the internet. (**Exposure**)	*A lover*	*A lover who works in technology* They know where to post the photos for maximum exposure and less chance of getting removed.

SOURCE: IAPP *Strategic Privacy by Design* (33), Copyright 2018 by R.J. Cronk, published by the IAPP

- So, look at the terms in parentheses in bold, like (**Blackmail, Exposure, Identification**). These refer to what we call ***privacy harms.***
- While these privacy harms can be subjective or objective, they can cause the recipient a lot of undue consternation. Sometimes they're legitimate, like law enforcement, but in many instances they're not.

CONCEPTUALIZATION OF PRIVACY

- It helps to create a conceptual diagram or flow to capture the breadth of activities that are encompassed by the concept of privacy.
- NYU Prof. Helen Nissenbaum's "contextual integrity" provides a privacy conceptualization:
  1. Privacy is provided by appropriate flows of information;
  2. Appropriate information flows are those that conform with contextual information norms;
  3. Contextual information norms refer to five independent parameters:
     a. Data subject
     b. Sender
     c. Recipient
     d. Information type
     e. Transmission principle
  4. Conceptions of privacy are based on ethical concerns that evolve over time.

ISSUES WITH CONCEPTUALIZATION (OF PRIVACY)

- While Prof. Nissenbaum's concept is encompassing, it is too abstract to clearly delineate a privacy violation from a non-privacy violation without reference to relevant laws, norms, etc.
- So, this can make things somewhat challenging for the BA and would warrant a **traceability matrix** to relevant policies and directives of the organization.
- Prof. Daniel Solove's empirical "A Taxonomy of Privacy" breaks privacy violations into four broad categories, based on his comprehensive survey of laws, precedent, and current norms.
- Three categories relate to information privacy during a. collection, b. processing, and c. dissemination and then a fourth category involves intrusion into seclusion as that data is possessed.

PRIVACY, SECURITY, & DATA GOVERNANCE

- Privacy is typically centered in the Legal Dept. of your org.
- Ergo, when putting on your BA hat and doing *Stakeholder Analysis*…you should put them near the top of your list!!
- This can be at times challenging for anyone with one foot in the "IT world", as we B.A.s tend to be… since the Legal Dept's privacy perspective tends to be one of laws and regulations which have a more narrow view of privacy impacts. Technology & its implications are an afterthought.
- You may find that communication between Legal and IT is infrequent, which is not uncommon.
- To wit: just think of how long it took legal safeguards to catch up with technologically-based ills like cyberbullying and "revenge porn".

PRIVACY VS. SECURITY

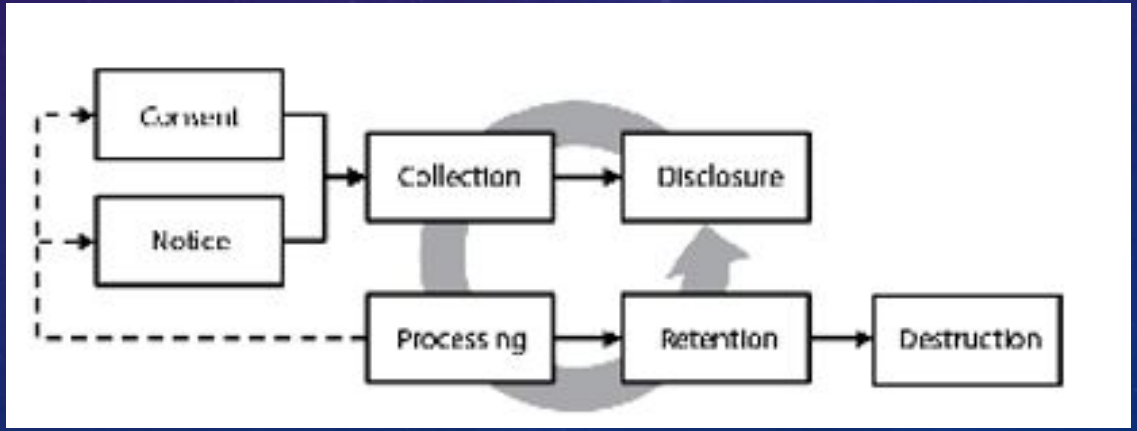
- When thinking about security, it's typically about the "CIA triad" – that is*, Confidentiality, Integrity*, and *Availability*.
- IN a sense, privacy is encompassing of security, and not the other way around; it's about the prevention of intrusions.
- But it also concerns a person's ability to control the *granularity* of information that others have access to. So it carries with it a certain degree of *autonomy*.
- However, the literature talks about "the privacy paradox", where users may claim that they wish their personal details to remain private, but in their actions they tend to reveal a great deal about themselves. In other words, users act contrary to their stated privacy preferences.

PRIVACY FRAMEWORKS AND PRINCIPLES: AN EVOLUTION

- *The Fair Information Practice Principles* (FIPPs) (1977), published by the U.S. Federal Trade Commission (FTC) and used as guidance to businesses in the United States.
- *The Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (1980), published by the Organization for Economic Cooperation and Development (OECD)
- *The Privacy Framework* (2005), published by the Asia-Pacific Economic Cooperation (APEC)
- *The Generally Accepted Privacy Principles* (GAPP) (2009), published by the American Institute of Certified Public Accountants (AICPA), and the Canadian Institute of Chartered Accountants (CICA)
- *NISTIR 8062, An Introduction to Privacy Engineering and Risk Management in Federal Systems* (2017), published by the U.S. National Institute of Standards and Technology.

THE DATA LIFE CYCLE RE: PRIVACY

- Your organization needs to specify upfront the purpose for which collected information will be used, and maintain consistency between actual uses and stated use.
- A **Privacy Notice**, presented upfront to prospective users of your web interface, is the primary vehicle to satisfy this need.
- Should there be any amendment to the purposes for which since-collected data is to be used, all users shall be notified promptly via email with an opt-in or opt-out provision as required.



SOURCE: *Introduction to Privacy for Technology Professionals* (22) v1.1, Travis D. Breaux, CIPT, Editor, Copyright 2020 IAPP.

COLLECTION METHODS

- Data Collection may occur at various points within a system; you can compose a data flow diagram or decomposed context diagram to elicit these.
- Data collection vehicles may include:
  1. **First-party collection**, e.g. the user submits a form with their data to the primary organization;
  2. **Surveillance**, when the collector observes data streams produced by the data subject without unduly interfering with the subject's behaviour;
  3. **Repurposing**, which happens when the previously obtained data is now assigned to be used for another purpose that was not in the original Notice, and
  4. **Third-party collection**, when previously collected data is transferred to a third-party to enable a new data collection flow.

TRADE-OFFS OF BUSINESS VALUE MAX. AND THE PRIVACY IMPERATIVE

- The data lifecycle is shaped by the privacy objectives and business practices of an organization, and the systems that they develop must be calibrated and adapted to satisfy these objectives.
- Basically, there are two opposite perspectives:
  1. A *maximize-information-utility* objective, which views data as the basis for monetization and a potentially increased stream of revenue, hence the impetus to collect as much data as possible, and
  2. A minimize-privacy-risk objective, which views data as potentially toxic with inherent risks that can result in egregious privacy harms.

TABLE: THE DATA-PRIVACY LIFECYCLE WITH TRADE-OFF PERSPECTIVES

	Maximize Information Utility	Minimize Privacy Risk
Collection	Collect any data that is available, as the value will be realized later when we envision new services and products; post generic privacy notices to accommodate broadly defined, future collections	Only collect data for established purposes and always collect consent from data subjects for sensitive data; allow data subjects to opt out of services they deem unnecessary and before collecting the data, when possible
Processing	Ensure open access to data within the organization; easy access drives innovation and creative new uses lead to increased utility and market competitiveness	Only use data for the purpose of the original collection; any new uses require additional consent from the data subject, and/or the sending of new privacy notices
Disclosure	Enable disclosures with third parties to leverage new marketing and outsourcing opportunities or to enable previously unplanned third-party services	Limit disclosures to those purposes for which data was originally collected; any new disclosures require additional consent from the data subject, and/or the sending of new privacy notices
Retention	Retain data as long as reasonably practical; long-term retention enables longitudinal analysis of data subjects to better accommodate their long-term needs and to build lifetime services	Destroy data when it is no longer needed to complete the transaction; any new uses that motivate longer retention periods require additional consent from the data subject and/or the sending of new privacy notices
Destruction	Avoid destruction by using long-term backups, or reduce access to data, but retain original data or a summary of the data for future uses or for reinstating services	As soon as data is no longer needed, ensure the data and any derivatives are removed from all systems using appropriate methods to prevent recovery

SOURCE: *Introduction to Privacy for Technology Professionals* (25) v1.1, Travis D. Breaux, CIPT, Editor, Copyright 2020 IAPP.

AGILE METHODOLOGY - CONSIDERATIONS

- Boehm & Turner are software engineering experts (who you may have heard of). Boehm discovered that the cost to fix a requirements defect during implementation increases a hundredfold, as compared to during the requirements phase; ergo the need to “bake in” privacy requirements.
- As Boehm and Turner argue, even Agile methods can include some planning when the risks outweigh the benefits of increased agility.
- As many of us know, in the Scrum process, the P.O. or Product Owner sets the priority on requirements – which are called **user stories** – and these are developed during each iteration (“sprint”).
- To bridge any perceived gap between Scrum and privacy, whoever is doing B.A. work ought to refine these user stories to help ID privacy risks and harms, then propose ways of reducing those risks.

DOCUMENTING REQUIREMENTS

- We must continue to distinguish between functional and *non-functional* requirements, even in the privacy realm.
- Examples of functional privacy requirements:
  - "The system shall provide a link to a privacy notice at the bottom of every page."
  - "The system shall obscure credit card numbers on the screen with * tokens, using AES 256-bit encryption."
- Examples of non-functional requirements:
  - "The system shall not disclose personal information without authorization or consent."
  - "The system shall clearly communicate any privacy preferences to the data subject."
- We tend to think of privacy requirements as being more associated with the latter group, since, after all, they don't directly provide any "pure" functionality to the user other than being a pre-requisite to the functionality at hand.

LOGGING (THE “PAPER TRAIL”)

- Sometimes, the auditor stakeholders [may be internal or external] associated with your project will demand you provide a LOG or “paper trail” of user interactions & consent, to satisfy them.
- For instance, you may be collecting site browsing metrics from users, in which case the user’s website history will be collected *but* will be linked to the fact of the user’s consent provided.
- Or, in other situations, you might opt for a requirement for a checkbox for more explicit consent, occurring at the start of a given transaction with at least two design hand-offs:
    1. Recording the date & time of the selection in a database, or
    2. Restricting the transaction to those who select the checkbox, in which case evidence of the collected data subsequent to the checkbox can exist *only if* the checkbox was ticked.

TECHNIQUES RE: PRIVACY REQ'TS.

- Using tables, matrices, and bullet points for privacy related requirements can also be supplemented by more visual material.
- Such as:
  - Process flow diagrams ("swimlanes")
  - Information or Data flow diagrams
  - Role & Permission matrices
  - State flow diagrams
- When it comes to elicitation, you may use:
  - Focus Groups
  - Interviews
  - Case Studies

DOCUMENT ANALYSIS & REQUIREMENT LINEAGE

- *Generally*, standards and guidelines are written to more easily infer requirements and be conducive to good solution design. Which also supports traceability / requirements lineage.
- However, when it comes to *less structured texts* – such as laws and regulations – a liaison with your LPR (legal, public & regulatory) stakeholder will almost certainly be required for interpretation.
- *Legal standards* refer to non-functional requirements that can span a conceptual solution design.
- Consider the following excerpt from the *COPPA Rule*: COPPA, 16 CFR Part 312

§312.5 Parental Consent.

(a) General requirements.

(1) An operator is required to obtain verifiable parental consent before any collection, use, and/or disclosure of personal information from children . . .

(b) Mechanisms for verifiable parental consent.

(1) An operator must make reasonable efforts to obtain verifiable parental consent . . .

(2) Existing methods to obtain verifiable parental consent that satisfy the requirements of this paragraph include:

USING TRACEABILITY MATRICES

- Thinking of *verbal notation* in entity relationships or data flow, we say that:
  - A trace link from a requirement to a privacy law means the requirement **implements** the law, whereas;
  - A trace link to a design element (e.g. RBAC mechanism) signifies that the requirement is implemented **by** the design element.
- Generic excerpt/example:
  - REQ-32: The System shall only disclose the minimum info required to complete a transaction.
  - REQ-35: The system shall provide law enforcement access to personal info by administrative subpoena.

Privacy Laws and Standards	REQ-32	REQ-33	REQ-34	REQ-35	. . .
Data Minimization					
HIPAA Privacy Rule, 45 CFR §164.502(b), Minimum Necessary Standard	X				
OECD Use Limitation Principle	X	X			
GAPP 5.2.1 Use of personal information		X	X	X	
Government Access					
Stored Communications Act, 18 U.S.C. §2703 (b)(1)(B)(i)				X	
GLBA, 16 CFR §313.15(a)(4)				X	

SOURCE: *Introduction to Privacy for Technology Professionals* (57) v1.1, Travis D. Breaux, CIPT, Editor, Copyright 2020 IAPP.

GOAL-BASED ANALYSIS

- **Goal-based analysis** can be applied to privacy policies to identify ***protections***, which are statements intended to protect a user's privacy, and ***vulnerabilities*** (which threaten a user's privacy).
- Take a look at this excerpt from Google's privacy policy:
  - WE may collect device-specific information (such as your hardware model, OS version, unique device identifiers, and mobile network information including phone number). Google may associate your device identifiers or phone number with your Google Account. [VULNERABILITY]
  - *We encrypt many of our services using SSL.* [PROTECTION]
  - *We offer you two step verification when you access your Google Account, and a Safe Browsing feature in Google.* [PROTECTION]

PRIVACY-ENHANCING CONSIDERATIONS

- These are considerations that persist throughout the data lifecycle:
  - Remove or generalize preconditions, e.g. such as retention and disposal of data from another province or state, not just the province or state in which the organization is based. Oftentimes, preconditions aren't warranted.
  - Preclude preconditions, assume exceptions, e.g. if a breach notification is contingent upon unencrypted data, then set a requirement to encrypt users' data to assume the exception and preclude that precondition.
  - Refine by refrainment, e.g. to meet a gov't requirement/rule to avoid improper disclosure, which does not otherwise state that stealing a cryptographic key can lead to disclosure, you can impose role-based restrictions on your users to prevent access to such a key – which ultimately supports the rule.
  - Reveal the regulatory goal – by identifying the regulatory goal and expanding it in broader terms to a variety of solution components and systems, IT developers are less likely to assume exceptions and construe the requirement more narrowly.

ANTI-GOALS (“MISUSE CASES”)

- Anti-Goals are an attacker’s own goals or malicious obstacles to a system.
- So, the way goal-oriented analysis works is for the BA in conjunct w/ the privacy engineer to identify the solution’s positive goals, then analyze the “inverse” of that: what would a malicious attacker do and how might it curtail the solution’s ability to achieve its positive goals (which are privacy-centric).
- You can model anti-goals as a hierarchy of sorts, whereby the blocks represent goals and anti-goals, and the arrows point from sub-goals to achieve higher-level goals.
    - E.G. a corrupt nurse steals a celebrity patient’s record (the sub-goal), then sells it to tabloids for his/her personal gain.

APPROACHES TO ANONYMIZATION

- The simplest approach is *suppression*: removing identifying values from a record. Names and identifying numbers typically undergo suppression.
- Another common approach is *generalization*: this is done by replacing a more specific element with a more general one, e.g. removing the day and month from a DOB or removing the year from display.
  - This arrangement can then be used for another requirement for user identification i.e. put in the day and month of your birthday… but you would not typically ask for just the year.
- A third approach is *noise addition* – this is the foundation of differential privacy. See next slide

DIFFERENTIAL PRIVACY

- **Differential privacy** is aimed at increasing privacy while upholding a modicum of accuracy of the underlying data.
- It deliberately injects “noise” in the dataset to perturb its characteristics, making identification of specific users less probable / more cumbersome.
- Some cloud platforms, like MS Azure, employ differential privacy as a feature.
- It is uses an Epsilon (ε) parameter, which typically ranges from 0.0 to 1.0, and where a value closer to zero means greater privacy but less accuracy, and vice-versa.

TRANSPARENCY, USER RIGHTS, AND DECISIONAL INTERFERENCE

- Transparency and user rights are fundamental concepts of privacy legislation and guidelines world-wide, and are found in organizations like the OECD, the U.S. FTC., and the EU's GDPR.
- Specific requirements can vary, but any organizations that collects or process PII [Personally Identifiable Information] must be transparent about their data practices AND to inform data subjects about their rights and responsibilities when it comes to certain data practices.
- The above is known as the notice and choice model.

PRIVACY CAVEATS: AN ANECDOTE

- The **“privacy paradox”** has been covered in behavioural science studies, and it comes down to the observation that even though many people EXPRESS concern over sharing their personal info, those same people have no qualms in providing or posting it to a given online presence.
- In those cases, complaints about unexpected discoveries or use of their data are not uncommon!
- I believe that a lot of this has to do with “FOMO”, or fear of missing out. As social creatures, we want to be part of certain shared experiences; there’s a hedonistic motive behind it.
- Almost needless to say, this puts certain obligations on you as an “ethical BA” – where are your user’s blind spots, and how can you arrive at a system that doesn’t unduly entice them to their detriment?

TECHNOLOGY CHALLENGES & EVOLUTION

- When it comes to privacy law & policy, you can be faced with the dilemma of potentially conflicting policy instruments.
- And a lot of the time, this is due to the lagged effect of technology nuances. Remember the earlier points on how the law tended to lag behind technological trends.
- So, developing tailored engineering solutions for one law may be cost-prohibitive or cumbersome, OR it may outright contradict the requirements of another.
- At this point it bears mentioning that there are two Chinese words that are inversions of each other, “wei-ji” meaning crisis, and “ji-wei” meaning opportunity!

TURNING TO THE DARK SIDE: DECISIONAL INTERFERENCE

- In some regrettable and even deplorable cases, some companies may use manipulative techniques to nudge people to relinquish their privacy. SUCH AS:
  - **Privacy Zuckering** – makes the process of changing the default privacy settings cumbersome for a typical end-user, by arranging a bad presentation of the available settings.
  - **Bad Defaults** – sometimes default options are deliberately chosen badly to encourage the [extra] sharing of personal data.
  - **Forced Registrations** – when a user is forced to register for an account to use part of the features or functions of a service, yet this is technically unnecessary.
  - **Hidden Stipulations** – terms & conditions are often difficult to read for the average end-user without any legal knowledge, which makes it possible to hide malicious stipulations.

STRATEGIC CATEGORIES FOR DATA PRIVACY

- The first four categories under ‘Strategy’ are data-oriented, and tend to be inward-facing.
- The next four categories are outward-facing and are more process-oriented.
- The final two, Inform and Control, are directed at the individual user.

Strategy	Dark Strategy
Minimize	Maximize
Separate	Centralize
Hide	Publish
Aggregate	Preserve
Enforce	Violate
Demonstrate	Fake
Inform	Obscure
Control	Deny

DATA-ORIENTED STRATEGIES, IN BRIEF

- To **separate** user data means to place it in either different physical or logical locations, in which the latter is typically restricted by user role groups or permissions.
- To **minimize** user data occurs at the collection stage, when the company only collects the bare minimum required to conduct operations of their online solution and business model.
- To **hide** user data is to "block out" or replace revealing characters with non-revealing ones, basically to obscure it from view. Like separation, this can be based on granular permissions.
- **Aggregating** user data typically means summarizing it to get macro measures like mean or standard deviations; or it might involve adding "noise" to it, to perturb it in some way so that its aggregate properties remain intact, but ID of individual users is made much less likely.

PROCESS-ORIENTED STRATEGIES, IN BRIEF

- To **enforce** data privacy is to apply administrative policies on proper data use, which may include physical security, when to use encryption, and context sensitivity.
- **Demonstrating** data privacy would show that you are processing personal data in a privacy-friendly manner, and may include logging, auditing, and reporting as proof.
- To **inform** in regards to data privacy is to overcome information asymmetries between the org and the user; it keeps users in the know on any data processing, or notifies them promptly in the case of breaches.
- To apply **control** w.r.t. data privacy is to provide data subjects with the means to process their own personal data; it basically “returns a degree of user autonomy”.

CONSEQUENCES OF PRIVACY VIOLATIONS: OBJECTIVE VS. SUBJECTIVE HARMS

- Privacy expert Ryan Calo splits privacy harms into subjective and objective harms.
- Calo defines objective harm as “forced or unanticipated use of personal information”.
- Personal information volunteered for use, cannot, by definition, result in a privacy harm; but forceful and non-consensual use of that info *can*.
- Objective harms may be loss of a job (or offer), divorce, or financial loss; subjective harms might be embarrassment or shame in some form.

DISPLAY OF PRIVACY HARMS BY CATEGORY

- Whenever there's a change in use by the recipient of the data from the original intent, or an asymmetry in understanding between the parties, a privacy harm is deemed to occur.
- Basically, it comes down to information imbalances, from a fairness and reasonableness perspective.
- Whether the individual really understood what they were providing their data for can be open to debate, however.

Table M: Privacy Violations, Harms and Adverse Consequences

	Information Collection	*Information Processing*	*Information Dissemination*	*Invasion*
Subjective Harms	**Psychological** (Embarrassment, anxiety, suicide[8])			
	Behavioral (Changed behavior, reclusion)			
Objective Harms		**Lost Opportunity** (Employment, insurance, housing, education)		
		Economic Loss (Inconvenience, financial costs)		
		Lost Liberty[9] (Bodily injury, incarceration, death)		
		Social Detriment (Lost of trust, shunning, ostracism, banishment)		

SOURCE: *Introduction to Privacy for Technology Professionals* v1.1 (123) Travis D. Breaux, CIPT, Editor, Copyright 2020 IAPP.

THE IMPERATIVE TO "BAKE IN" PRIVACY

- Again, I re-emphasize the need to always adhere to Dr. Cavoukians 7 principles of privacy that I covered at the outset.
- As another "magic 7" bit of wisdom… consider Stephen Covey's 2nd habit of highly effective people: "BEGIN WITH THE END IN MIND".
- We need to differentiate between Compliance by Design vs. Privacy by Design; the latter requires a properly defined and understood goal at the outset, not *post facto* analysis…which is the former design type. That is more process-oriented and reactive.

QUALITY ATTRIBUTES RELATING TO PRIVACY

- Besides meeting the purpose at face value, you may have service integrity requirements; these provide proof that a domain actor has legitimately performed some activity within acceptable privacy boundaries.
- This is essential if you have some sort of Auditor stakeholder, who will need to periodically sample proof that you have “baked in” privacy requirements from the get-go.
- Such quality attributes can be “two birds for one stone”…satisfying business need for justifying revenue or establishing contract for accrual of $$, as well as capturing of privacy grounding to fulfill “CYA”.

TIME FOR QUESTIONS…

AND COMMENTS AS TO PERCEIVED UTILITY IN YOUR OWN WORK!