AGENDA

• Data Evolution
• Error Breaches
• Privacy-by-Design
• The Power of Defaults
"Between the dawn of civilization and 2003, we only created five exabytes of data; now we're creating that amount every two days. By 2020, that figure is predicted to sit at 53 zettabytes (53 trillion gigabytes, the equivalent of about half a billion HD movie downloads) – an increase of 50 times."

*Hal Varian, Chief Economist, Google*

- **Data = Value** and with a treasure trove of data available, privacy breaches are becoming more prevalent and far-reaching:
  - **Facebook / Cambridge Analytica**: 87 million profiles transferred and analyzed for behavioral patterns and targeted for political persuasion.
  - **Equifax**: 148 million records exposed containing personally identifiable information (PII), to include credit data, driver’s license numbers, social security numbers (SSN), dates of birth, phone numbers, and email addresses.
  - **Office of Personnel Management (OPM)**: 25 million compromised files on current and former federal government employees included personal information, background checks, fingerprints, and adjudications for security clearances. OPM’s director and Chief Information Officer resigned.
Lack of data awareness (didn’t know PII was included)

Data Loss Prevention (DLP) tools can only do so much (lexicons have limitations)

Profile building of individuals (data commingling can potentially re-identify people or build out profiles)
California Attorney General released a report in 2016 on data breaches from 2012 to 2015.

California law requires reporting data breaches to the Attorney General.

Government sector data losses were primarily due to employee error.

What are error breaches?
Errors by insiders (i.e., leadership, employees, contractors) that resulted in breaches included sending information by email to unintended people and accidentally exposing information to unauthorized people (e.g., screen views or online posts including text or images).
Privacy by Design (PbD) calls for privacy to be taken into account throughout the entire engineering process.

- Privacy by design integrates privacy considerations throughout the information life cycle.
- It helps to make sure that people and organizations are asking the right questions so that technology solutions have privacy protections embedded upfront.
- The Software Development Life Cycle (SDLC) can include privacy gates for software and application development projects.

What is the County of Santa Clara doing?
- Established an Enterprise Data Risk Review Committee & Working Group
- Will work to integrate privacy gates with the SDLC and other technology and County processes (e.g., Privacy Vendor Vetting Checklist as part of the Procurement RFP process)
- Conducting Privacy Impact Assessments (PIA) and Sensitive Information Inventories for County information systems.

PbD can be integrated with the SDLC process through privacy gates or checkpoints to make sure privacy is baked into solutions rather than bolted on post-deployment.
We now carry and interact with devices that collect thousands of data points continuously.

On their own or combined, this data can reveal extremely detailed (but never complete) information about our lives: where we live, work, and play, who we spend time with, what we buy, our health status, our hopes and fears, our financial security (or lack of it).

Collection of this data is often on by default.

Organizations that build software and the people who use it often do not change these defaults, for many reasons.

- People don’t know about them
- People don’t understand them
- Choice fatigue
- Psychological short-cuts

Government entities should think about the relationships, requirements, and controls they have on their constituent data when designing systems, working with third parties, and data sharing by default.
What are real-world examples?

**Example – Show me the money**

- Many retirement plans are **opt-in**, meaning you need to **actively** request HR to enroll you.
- Research has shown that only around **60%** of people enroll or **opt-in** to retirement plans.
- Making investment decisions can seem overwhelming to people and defaults assist in this way, allowing for automated decisions to be made in your “best interest.”
- More recently, workers have been automatically enrolled in retirement plans (**opt-out**) rather than left to opt-in themselves. This spikes the numbers up to over **90%** enrollment.

**Example – Defaults in health**

- In the United States, the default is **NOT** to be an organ donor, that is you must **actively** check a box on the DMV questionnaire to **opt-in** to the organ donation program.
- However, in most European countries, the default **IS** to be an organ donor. You must **actively** uncheck a box to **opt-out** of the organ donation program.
- What effect do these defaults have? Opt-out countries have much higher donation rates than opt-in countries:
  - 90% participation rate in opt-out countries
  - 15% participation rate in opt-in countries
Technology departments, privacy offices, and design experts can work together to protect data, apply privacy-by-design, and set defaults.

• With the expanding use of data, structured and unstructured, IT professionals should consider the need to balance the need to share data with the need to protect it.

• Integrate privacy as part of the planning, procurement, design, and development processes.

• Work with Privacy Pros to understand risks upfront, which will help to prevent breaches and privacy violations later.

• Additionally, knowing the impact of defaults are part of a larger system of choice architecture, which is the design of different ways in which choices can be presented to consumers or constituents.

• When designing apps & systems, consider defaults and how they should be set, to include:
  • Options in menus
  • Coding applications
  • Presenting information in a specific format or order
  • Opt-in over opt-out
  • Applying configuration settings in a smart way
  • ... or put the donuts in a really inconvenient place.