Introduction

The world is certainly flat. Everyone said so. The government said so. The church said so. Your wise old aunt and the richest guy in town said so. Everyone.

Except, a few explorers, dreamers, scientists, artists, and plain-spoken folks who looked out at a sky that looked more like a bowl and noticed that the ground and sky always met for a brief kiss before the observer wandered ever closer and the meeting became elusive. And shadows, tides, and other indications seemed to suggest that there might be something more than dragons beyond the edge of the world. And so, as it turned out, the world was not, in fact flat. There was a seemingly endless set of new possibilities to discover.

Privacy is *certainly* dead. Everyone said so. Rich people with big boats who sold stuff to the CIA in the 1970s said so. Founders of important hardware companies said so. Someone who blogs said so. The government cannot make up its mind which person should say so or if the polling numbers look right, but it might say so. Someone tweeted. Even really old technologists who helped invent the whole thing said so. *Everyone*.

Except, a few explorers and inventors and philosophers and children and parents and even government regulators who looked out at a seemingly endless sea of data and could still see how a person can be distinguished from a pile of metadata. This is true for people who wish to decide for themselves the story they wish to tell about themselves and see a different horizon. The privacy engineer sees this horizon where privacy and security combine to create value as a similarly challenging and exciting time for exploration, innovation, and creation; not defeat.

The purpose of this book is to provide, for data and privacy practitioners (and their management and support personnel), a systematic engineering approach to develop privacy policies based on enterprise goals and appropriate government regulations. Privacy procedures, standards, guidelines, best practices, privacy rules, and privacy mechanisms can then be designed and implemented according to a system's engineering set of methodologies, models, and patterns that are well known and well regarded but are also presented in a creative way. A proposed quality assurance checklist methodology and possible value models are described. But why bother?

The debate about data privacy, ownership, and reputation poses an irresistible and largely intractable set of questions. Since the beginning of recorded history, people have sought connection, culture, and commerce resulting from sharing aspects about themselves with others. New means of communication, travel, business, and every other social combination continue to evolve to drive greater and greater opportunities for the solo self to be expressed and to express oneself in person and remotely. It is all terribly exciting. Yet, every individual desires a sense of individuality and space from his or her fellow man; a right to be left alone without undue interference and to lead his or her individual life.

Governments have played a stark role in the development of data privacy. Laws are created to protect, but there are also abuses and challenges to individual rights and freedoms in the context of multiple governments in a world where people have become free to travel with relative ease and comfort—sans peanuts—around the globe and back again. National and international security norms have been challenged in both heroic and embarrassing fits and starts. The role of total information vs. insight and actionable information is debated again and again. "Insiders" and fame seekers have exposed massive data collection programs.

In the information technology sector, data privacy remains a matter for heated debate. At times the debate seems as if technologists somehow wished (or believed) they could escape the norms of general social, cultural, and legal discourse simply by designing ever more complex systems and protocols that "need" increasing levels of sensitive information to work. The lawyers come trooping in and write similarly complex terms and conditions and hope to paper over the problem or find a cozy loophole in unholy legislative agendas. Investors search in vain for beans to count. Everyone else finds privacy *boring* until their own self-interests are compromised.

At the same time, just as automotive technology eventually became a ubiquitous and necessary part of many more lives, so too has information technology, from phones to clouds, become such an essential part of industrialized nation-states. Personal data fuel and preserve the value of this new information boom. Thus, the technical elite no longer can dismiss the debate or pretend that data privacy doesn't matter, nor can they fail to build new creations that defy basic privacy precepts, which we will discuss herein, if they wish to see this new world unfold and grow.

If an executive at a global company publicly were to state that he doesn't believe in taxes and therefore will not pay them to any government, he would likely be removed or at least considered to be a great humorist. Not so for data privacy in the past. In the past decades, executives and other makers and consumers of information technologies regarded data privacy as some sort of religion that they could believe in or not at will and without earthly consequence. They certainly did not regard privacy as a *requirement* to measure, to debate in the boardroom, or to build at the workbench. We see these uninformed days of privacy as religion as nearly over. The age of data privacy as a set of design objects, requirements for engineering and quality measures, is dawning, and we hope to help the sun come shining in.

In fact, plain old-fashioned greed and an instinct for value creation will *hasten* the age of privacy engineering and quality. We know that the concept of privacy regarding one's person, reputation, and, ultimately, what can be known about the person has been the inspiration of law and policy on one hand, but we also know that innovation and the recognition that privacy—informational or physical—has value.

Andrew Grove, cofounder and former CEO of Intel Corporation, offered his thoughts on Internet privacy in an interview in 2000:

> *Privacy is one of the biggest problems in this new electronic age. At the heart of the Internet culture is a force that wants to find out everything about you. And once it has found out everything about you and two hundred million others, that's a very valuable asset, and people will be tempted to trade and do commerce with that asset. This wasn't the information that people were thinking of when they called this the information age.*[4]

Thus, people living in the Information Age are faced with a dichotomy. They wish to be connected on a series of global, interconnected networks but they also wish to protect their privacy and to be left alone—sometimes. Both business and governmental enterprises, striving to provide a broad base of services to their user community, must ensure that personal information and confidential data related to it are protected. Those who create those systems with elegance, efficiency, and measurable components will profit and proliferate. History is on our side.

We call the book and our approach "privacy engineering" in recognition that the techniques used to design and build other types of purposefully architected systems can and should be deployed to build or repair systems that manage data related to human beings.

We could have similarly called the book "design principles for privacy" as the techniques and inspirations embraced by the design communities in informatics, critical design, and, of course, systems design are also a part of the basic premise where one can review an existing successful framework or standard and find inspiration and structure for building and innovation. The very nomenclature known as privacy engineering is left open to the possibility of further design.

The models shown are abstractions. Models are never the reality, but models and patterns help designers, stakeholders, and developers to better communicate and understand required reality.

Confidence in privacy protection will encourage trust that information collected from system users will be used correctly. This confidence will encourage investment in the enterprise and, in the case of charity enterprises, will encourage people to donate.

There are many books and papers on privacy. Some focus on privacy law, others on general privacy concepts. Some explain organizational or management techniques. This book is intended to be additive. This book crosses the boundaries of law, hardware design, software, architecture, and design (critical, aesthetic, and functional). This book challenges and teases philosophical debates but does not purport to solve or dissolve any of them. It discusses how to develop good functionalized privacy policies and shows recognized methodologies and modeling approaches adapted to solve privacy problems. We introduce creative privacy models and design approaches that are not technology specific nor jurisdiction specific. Our approach is adaptable to various technologies in various jurisdictions.

---

[4]"What I've Learned: Andy Grove," *Esquire*, May 1, 2000.

Simply put, this is a book of TinkerToy-like components[5] for those who would tinker, design, innovate, and create systems and functional interfaces that enhance data privacy with a sustainability that invites transparency and further innovation. We wish to demystify privacy laws and regulations and nuanced privacy concepts into concrete things that can be configured with flexible, engineered solutions.

The *Privacy Engineer's Manifesto: Getting from Policy to Code to QA to Value* is a unique book. We introduce privacy engineering as a discrete discipline or field of inquiry, and innovation may be defined as using engineering principles and processes to build controls and measures into processes, systems, components, and products that enable the authorized processing of personal information. We take you through developing privacy policy to system design and implementation to QA testing and privacy impact assessment and, finally, throughout the book, discussions on value.

- Chapter 1 discusses the evolution of information technology and the network and its impact on privacy.
- Chapter 2 discusses a series of definitions: policy, privacy engineering, personal information (PI), and the Fair Information Processing Principles (FIPPS).
- Chapter 3 covers data and privacy governance, including data governance, Generally Accepted Privacy Principles (GAPP), Privacy by Design (PbD), and other governance frameworks.
- Chapter 4 introduces a privacy engineering development structure, beginning with the enterprise goals and objectives, including privacy *objectives*, that are used to development privacy policy.
- Chapter 5 discusses privacy engineering requirements. We then introduce use cases and use-case metadata.
- Chapter 6 introduces enterprise architecture and the various views of it. We dig into the privacy engineering system engineering lifecycle methodology. We show the Unified Modeling Language (UML) usage flow from the context diagram, using the UML use-case diagram, to the use of business activity diagrams, including showing key data attributes, then on to data and class modeling using the UML class modeling diagram, and then to user interface design. We use the system activity diagram to show where FIPPS/GAPP requirements are satisfied within the privacy component design (scenario 1) and then we move to dynamic modeling where we define service components and supporting metadata, including the inclusion of privacy enabling technologies (PETs). We then discuss the completion of development, the development of test cases, and the system rollout.

[5]See www.retrothing.com/2006/12/the_tinkertoy_c.html for a random, cool TinkerToy creation by MIT students.

- Chapter 7 discusses the privacy component app, which will be used to maintain the Privacy Notice. The privacy team, along with the data stewards, will enter and maintain the privacy rules. When an embedding program requires personal information, the privacy component will ensure that the personal information is collected according to privacy policies.
- Chapter 8 presents, as an example, a small mobile app, using a simplified version of the privacy component to support a high school cross-country runners app.
- Chapter 9 covers an example vacation planner app that utilizes a privacy component that has already been developed, tested, and implemented by a large hospitality company that requires a system to help its customer community plan a vacation at one of their hospitality sites.
- Chapter 10 covers quality assurance throughout the development lifecycle, data quality, and privacy impact assessments (PIA).
- Chapter 11 discusses privacy awareness assessments and operational readiness planning.
- Chapter 12 covers the organizational aspects of privacy engineering and aligning a privacy function to IT, to data governance or data stewardship, and to the security management function.
- Chapter 13 discusses how data and data privacy may be valued.
- Chapter 14 covers our musings about the future of privacy and privacy engineering along with our Privacy Manifesto.

Why Anyone Should Care About Privacy, Privacy Engineering or Data at All

It's time to serve humanity.

> Humanity is people.
> Humanity is empowered stewardship of our surroundings—
> Our universe, planet, and future.
>
> Humanity is described by data;
> Data about humans;
> Data about all things human.
>
> Data is not humanity;
> Data tells a story;
>
> Data is leverage;
> Data is not power.

Humanity can capture data.
Data cannot capture humanity.

It's time to serve humanity.
There is no one else.
We are already past due.

This is the paradox in which the privacy engineer discovers, inspires, and innovates.
Let's begin.

PART 4

Where Do We Go from Here?

CHAPTER 13

Value and Metrics for Data Assets

It is the mark of an educated mind to rest satisfied with the degree of precision which the nature of the subject admits and not to seek exactness where only an approximation is possible

—Aristotle

Or, put another way, don't go over thinking things—or over measuring things.

—Steve Weiss, Editor

No precision is possible to quantify or qualify the value of data, well or poorly designed system efficiencies, or brand value if we fail to begin. Yet, the reality is that enterprises run on well-trod resources such as money, real estate, and property. They also run on brand loyalty, percentage of churn, customer satisfaction, and leverage. The point here is that it is hard to measure the value of intellectual or virtual property such as the right to use, process, or remain a fiduciary for data. This chapter will put forth some ideas and concepts about potential data or data-centric systems. A privacy engineer holding this book will recognize that, here too, is a topic rife with opportunity for quiet incremental improvement and bold innovation.

One of the most elusive, yet impactful, tasks before the privacy engineer is to find measurements for incremental progress in designing and executing data governance standards and utilities and to report those metrics in terms of value. Value may come in many forms:

- Qualitative value as in improved efficacy of data system flows and customer satisfaction
- Quantitative value in terms of:
    - Loss avoidance
    - Incremental gains in information-based products and services or those accelerated by PI
    - A lower percentage of churn
    - Lower perceived "creepiness"

It makes sense here to have a little refresher from a discussion we began in Chapter 2 that covered some of the differences among privacy, confidentiality, and security before addressing value and metrics directly. These differences are particularly interesting, as data privacy tools and models are built, differentiated, and measured for value creation among a thicket of security or general "compliance" goods and services.

Data privacy is, in a very real sense, the most immature of the categories of intellectual property (IP), even though its roots travel far back in time. Traditional notions of IP include patents, trademarks, copyright, trade dress, trade secrets, and the contractual or social concepts of confidentiality. Of course, these notions often offer up models of "ownership" or "control" beyond that comfortably conceived for data privacy and protecting information about humans, but the models are helpful when discussing or determining measurement or quantitative models deployed to arbitrarily value it.

Trademarks (and other IP analogous legal objects) designate the origin of a good or service. For better or worse, a trademark's social utility is to alert end users to the origin or owners, creators, or controllers of goods or services. As part of the exchange for a limited monopoly right to trade goods under an exclusive mark, the owner of the trademark has a bundle of rights and obligations (assets and liabilities) associated with such ownership. For example, under US law,[1] a trademark owner must police his mark to be sure consumers are not fooled into believing imposters' goods are masquerading as his own (the cost of these efforts may be viewed as an expense undertaken for securing or protecting the right to remain the sole source of goods). Similarly, an IP owner must also ensure that goods or services are of a consistent quality (another cost center or liability undertaken both to protect the asset and protect the consumer). On the balanced side of the economic valuation, a trademark owner is entitled to have a limited monopoly as the source of a good or service as a direct market advantage and is also entitled to gain an extra boost and intangible advantage as a greater brand strategy to build emotional or other customer equity.

Data privacy may be considered as the bundle of rights and obligations that arise from the data emanating from or describing a person. Whereas the trademark owner is the origin of the good or service, so too is the human an identifiable individual data subject the origin of personal information. Current laws, regulations, and culture create the obligations for those who wish to remain fiduciaries or processors of data, and those same contextual requirements also create a platform for opportunities for asset management and leverage.

There are a number of imperfect analogies and models to help guide the way to begin the measurement and evaluation of the asset and liability balance for data privacy. None are perfect, but they are a good start in the absence of existing practices. (Remember Aristotle: don't seek exactness when only approximation is possible.)

[1]In other countries, laws around IP differ much as they do for data protection as a reflection of local or regional custom and commerce. A trademark owner may, for example, be allowed to own a trademark for a certain period of time without proving commercial use of that mark or have differing rights in his ability to alienate his rights to the mark.

DO WE TREAT DATA AS ASSETS?

By Rena Mears, Managing Principal of RMCS, LLC

"We treat data as an asset . . ."

A ubiquitous phrase found in hundreds of thousands of online privacy policies[2] that succinctly conveys a sense of shared value and due care on the part of the enterprise to the web site user. Given its widespread use in privacy policies, it may be surprising to note that managing personal information as an asset is still in the very early stages of development within most enterprises. Many of the basic asset management processes such as inventorying, cost analysis, and asset valuation are in a nascent state, and consequently the tools and processes considered standard when managing other enterprise assets may be nonexistent or only minimally applied to personal information (PI) assets. So is it worth the effort and cost to develop these processes? Does adopting a more asset-based approach support or inhibit the effective and efficient management of personally identifiable information in the enterprise?

To answer that question, it is important to consider the definition of an asset, the various uses of PI in the organization, and the impact of valuation on the allocation of enterprise resources and shareholder value. The definition of an asset is deceptively simple:

- A resource controlled by an entity
- As a result of a past event
- From which future economic benefits are expected to flow to the entity[3]

However, when the criteria are applied to PI, the complexity of the management challenge becomes readily apparent. Diverse cultural, regulatory, and marketplace requirements have an enormous impact on defining and managing PI assets. Where, when, and how data is acquired ("past event") can determine what is considered a PI asset, how it can be used, and the level of control that must be exercised to effectively manage the asset throughout its lifecycle.

In response to this complexity, the general tendency has been to treat all PI assets as similar in nature and manage them on a tactical level as a cost-center issue. This approach often results in some or all of the following:

- PI asset management processes focus on risk reduction and cost minimization rather than asset optimization.

[2]Internet search results from "treat data as an asset" "privacy policies."
[3]International Accounting Standards Board. (2003). International financial reporting standards (IFRS's): Including international accounting standards (IAS's) and interpretations as at. London: International Accounting Standards Board. Elements of financial statements (IAS 1 article 10)

- Senior management involvement is limited to crisis response (e.g., breach, regulatory, enforcement action) or periodic reporting of risk (e.g., changing law, audit findings) and does not extend to consideration of strategies to maximize return on the PI assets.
- Managing PI assets defaults to the midmanagement layer of the organization and is treated primarily as a legal and compliance issue.
- PI assets are maintained in silos and management may be inconsistent and unaligned with company strategy.
- Enterprise resources (e.g., budget, human capital, technologies) are allocated evenly across all PI assets regardless of the value of individual assets, resulting in misallocation of resources, hidden costs, and unnecessary expense.
- Inventory of PI assets is incomplete or nonexistent, thereby limiting management's ability to evaluate, manage, and optimize the asset.

Changing market conditions are forcing a reexamination of this cost-based approach to managing PI assets. Companies that once considered themselves solely product oriented now see themselves as "information-driven" businesses that rely on data assets, including PI assets, to compete effectively in the marketplace. Innovative technologies and reduced storage costs support the acquisition and mining of vast amounts of data. The rapidly expanding definition and changing role of PI assets in current business models is driving the need for a more nuanced approach to evaluating and managing these assets.

A utility-based approach to asset management examines the "usefulness" or net contribution of individual or subclasses of PI assets to the value chain of an organization. The approach considers the various use cases of PI assets to identify future economic benefits (e.g., revenues, product enhancement), associated costs, and potential risks to determine net contribution values. Assets with similar use cases, characteristics, and values may be grouped into asset profiles that form the basis for asset optimization through strategy development and the application of customized management processes. It is important to note that asset optimization of PI assets is not the same as merely maximizing direct revenue from the use of personal information. There are many use cases for PI assets, and enterprise utility may relate to support activities and contributions through risk or cost reduction (e.g., meeting legal requirements, optimizing talent acquisition). Some advantages that may be expected when adopting a utility-based approach to PI asset management are:

- PI asset management approach focuses more broadly on asset optimization and considers opportunities and risks beyond legal and compliance requirements.
- Senior management involvement extends to the development of PI asset strategies and supports enterprise recognition of the strategic value of PI assets.

- Management of PI is appropriately positioned at all levels in the organization, resulting in more efficient use and effective control of the asset.
- Enterprise resources (e.g., budget, human capital, technologies) are allocated in a more "value-based" manner, thereby focusing expenditures on assets with the highest contribution to the enterprise value chain.
- Basic asset lifecycle processes (e.g., inventorying, cost analysis) are applied to PI assets and may result in identification of new management options (e.g., "build or buy," outsourcing).
- Underperforming assets can be identified and managed appropriately (e.g., retired or deleted, access/use limitation).

Many organizations consider it too costly and very difficult to adopt a utility-based approach to PI asset management. However, the cost of not adopting such an approach may mean that PI assets continue to be treated as "white noise" in the enterprise, widely distributed throughout the organization and relatively homogeneous in nature. That approach ignores the very essence of the definition of an asset and will likely ensure that PI continues to be a source of high risk, hidden cost, and unnecessary expense to the enterprise. Suboptimized assets whose risks and cost outweigh their contributions are more commonly known as liabilities.

Finding Values for Data

> *Some day, on the corporate balance sheet, there will be an entry which reads, "Information"; for in most cases, the information is more valuable than the hardware which processes it.*
>
> —Rear Admiral Grace Murray Hopper

Values for data protection measures have been based on survey and anecdotal evidence relating to reported data breaches. Such breach reporting is typically thrust upon an enterprise by prevailing data breach legislation, best practices relating to credit monitoring or other services, and legal or marketing expenses undertaken in response to the negative perceptions caused by such breaches.[4] Another method for measurement

[4]"Ponemon study shows the cost of a data breach continues to increase." www.ponemon.org/news-2/23

may be to analyze prior fines or other regulatory requirements, such as Federal Trade Commission Consent decrees requiring as much as 20 years' oversight by a third-party audit company or other self-reporting mechanism.[5]

These traditional methods for data valuation fall short of the hoped for objective in a few fundamental ways. First, they are retrospective and often based on internal process or insider bad action—often quite difficult for an enterprise to anticipate or prevent. The incident may have arisen from a criminal actor, such as a hacker, or from product vulnerability in an increasingly complex IT ecosystem.[6] Second, the cost of a failure is but one component of risk avoidance—inefficiency, uncurated data mismanagement and waste, and, most important, true data asset prospective value are rarely addressed and even more rarely managed as sources of proactive investment.

Uncurated data is data that is not assigned to, owned by, or governed through specific methodologies or specific responsibilities. In short, this is data that is not being actively processed or organized to add value to either the data subject or the enterprise. For example, special events and business conferences require a great deal of personal data to accept payment, organize meetings, arrange travel, and more. Some of that data remains and grows in value as it is leveraged to build relationships with participants and personalize goods or services while the same data poses a risk only if left neglected or unused for its intended purpose.

Some data loses its relevance and becomes a compliance liability or risk where the data directly related to ended events or meetings for logistics, for example, is no longer needed for any relevant conference-related purpose. Retaining irrelevant portions of collected materials (or information) costs an enterprise money, time, and other resource expenditures. Although hardware storage may seem inexpensive and the myth persists that retention of data past its original purpose may create a "what if" or potential asset value, such is rarely the case. In fact, an enterprise may not have the legal right to process uncurated data if the future purpose of processing is beyond the original purpose.[7]

A mental experiment is helpful here, where a CFO continues to pay to store and move a warehouse filled with notebooks and pencils. These office supplies may be useful for future meetings or for scratch paper if date embossed. Nonetheless, if no one understands where the warehouse is located, if it has doors or a lock, and the nature of the supplies, and if no one has any responsibility for the warehouse's content,

[5]There are many examples of FTC Consent decrees and Data Privacy Authority sanctions with a variety of financial or other equitable remedies. In many countries, sanctions are either fines or undertakings to alter activities. In the United States, most federal-level penalties also contain the obligation for an enterprise to pay for annual audits of the enterprise privacy compliance efforts. See Microsoft's consent decree settling allegations with the FTC that the company made false statements regarding its ability to provide privacy or security to its customers. www.ftc.gov/opa/2002/08/microsoft.shtm. See also France's Commission nationale de l'informatique et des libertés (CNIL) sanctions against Google and its specific requirements that it hopes to impose on Google for its processing of French PI. www.cnil.fr/english/news-and-events/news/article/google-failure-to-comply-before-deadline-set-in-the-enforcement-notice/

[6]See "Predicting the unpredictable: Detecting chaos in mathematical equations." www.mit.edu/newsoffice/1998/chaos.html

[7]See OECD Guidelines Purpose Principle, discussed in Chapter 2.

the enterprise must continue to pay for its management, realizing no further value and risking further losses by fire or workplace injury for movers or other unexpected problems. Just as the information ecosphere provides the potential for massive data stores and assets, so too does it create the very real possibility for waste, loss, and unplanned risk.

KNOWLEDGE GOVERNANCE

By Kenneth P. Mortensen, Chief Governance Officer at CVS Caremark

> *What is a system? A system is a network of interdependent components that work together to try to accomplish the aim of the system. A system must have an aim. Without an aim, there is no system. The aim of the system must be clear to everyone in the system. The aim must include plans for the future. The aim is a value judgment.*
>
> —Dr. W. Edwards Deming,
> *The New Economics for Industry, Government, Education*

In the age of "big data" and "advanced persistent threats," a privacy professional can no longer focus solely on developing and implementing the processes and procedures to drive information governance, but rather she needs to advance her organization through the optimization of risk while facilitating core management decision making in order to create real value. This is the new world of "knowledge governance."

In the past, an organization looked simply to corral its data into a warehouse so that it could be understood which datasets and which data elements provided operational leverage within the activities or functions of the organization—otherwise known as "data governance." By producing a common or uniform view into the organization's data, data governance allowed, for the first time, an understanding of which data fed the organization's activities or functions. Nevertheless, this was a single dimension view that lacked the ability to understand the utility of the data within those activities and functions. Without a view to the data utility, an organization flies blind to legal and regulatory compliance issues, such as with privacy and information security. Thus just having a common understanding or reference model for the data of an organization does not open up those data for use and disclosure without significant risk regarding privacy and security.

From that gap, the privacy profession promoted the concept of "information governance" that allows for the data to communicate information. In literal etymological terms, information means to give form to something. In business terms, the word focuses on the ability to transmit data by providing form to a message by casting it into a profile or pattern for communication (sharing). This means definitions for information can be grouped roughly into quantitative and qualitative categories.

The qualitative definitions focus on the criteria that add meaning to the message that is communicated. The quantitative definitions focus on measuring the quantity of information units or the strength of its transmission. But this alone did not address the risks inherent with data governance. The governance aspect at the information level comes from the effective and efficient management of information within organizations. Management is the process of getting activities completed efficiently and effectively through the enterprise. The goal (or function) of management is to get the best return on enterprise resources by getting things done efficiently.

There are four basic pillars to any management process: plan, organize, direct, and monitor. An organization must, through data governance programs, *plan* the path for information within any organization as well as address any external collection or disclosure. Next, the organization will need to *organize* not only the data, which gets the organization only as far as data governance, but also the uses and disclosure to discover the utility of the information. From those uses and disclosures, the organization can *direct* protections and safeguards so that the organization can not only use the information thoroughly, but also in a compliant manner. Last, the *monitoring* of the processes and procedures is crucial to ensure that governance works to drive continuous compliance.

At this point, many organizations put down their tools, convinced that they have full use of their information in a methodology that ensures compliance with needed privacy protections and necessary security safeguards.

Unfortunately, these organizations, while able to survive the enforcement environments because they operate in a compliant manner, cannot progress into having full enterprise understandings of what value they can extract from all the information. Legal compliance does not optimize risk to the organization; nor does this coordination of effort address more than one facet of risk. The organization must look to all functionalities of the organization to understand the impact of risks associated with the information resources. To move to the next level and attain "knowledge," the organization must address information and its management strategically. Strategic management of information across the organization addresses not only the need to optimize the risk to the organization, but by establishing all the information as a critical organizational (or, better put, *enterprise*) asset, if not the most critical asset, the organization can introduce effective efficiencies into the decision-making processes for management, enhancing the return on the investment in information. An organization needs wide-ranging processes to capture not only data protection, but also data compliance, which takes in the complexity and diversity of the risk and legal environments. Knowledge is the value form of information, just as information is the communicative form of data. To accomplish this objective, an organization must employ enterprise governance that addresses all aspects of information within the organization with processes and procedures to deconflict and reconcile priorities to ensure governance efficiency.

Once knowledge governance has been achieved, an organization can extract the value of core data and information. The organization's leadership will be guided by this knowledge in advancing the goals and objectives of the organization, or as Dr. Deming noted when addressing similar issues from a quality management aspect:

> *The prevailing style of management must undergo transformation. A system cannot understand itself. The transformation requires a view from outside. The aim . . . is to provide an outside view—a lens—that I call a system of profound knowledge. It provides a map of theory by which to understand the organizations that we work in.*[8]

Knowledge governance for data assets can only be enhanced by further exploring other metric and valuation models. As is true for other sections of **The Privacy Engineering Manifesto**, methodologies and processes have been undertaken to create useful valuations of difficult-to-measure tangible and intangible inputs and outcomes. Data privacy is neither the most unique problem in the world nor the least measurable. Nonetheless, to quote the late American novelist David Foster Wallace, sometimes "the most obvious, ubiquitous, important realities are often the ones that are hardest to see and talk about."[9] Once discovered, the language of value for data privacy may be the key to opening the door to more practical matters.

Valuation Models

The following potential models should be viewed as a sketch pad of sorts; a group of potential techniques and tactics for assigning values or making concrete the value for data and data-centric systems. As technologies become more deliberately designed for data protection and policies evolve to become both legally more efficient and compatible with requirements setting, so too should valuation models evolve.

Model 1

Find something to count and count it:

- Data breach, customer churn after direct enterprise activity, or other regionally relevant contextual activity (such as a significant breach or a news-making threat or economic instability that causes data or customer contacts to increase or decrease).
- Leverage the GAPP maturity model and gauge costs to move to a higher maturity model. Balance cost against brand valuation, data reliant programs, or marketing events to the percentage spent to acquire customers.

[8]W. Edwards Deming, The New Economics for Industry, Government, Education, Ch. 4 (1994)
[9]"This is Water", Commencement Speech to Kenyon College class of 2005 written by David Foster Wallace.

- Read 10K annual reports or other publicly available, legally binding documents to find data-critical programs such as expansions into new jurisdictions, outsourcing, or cloud shifting business models or determine the geographic mix of customer or employees who provide critical data to the enterprise. Make an educated or sample-based guess regarding the importance of employee or customer data access based on these disclosures.
- Estimate IT spent regarding data-centric systems, and measure the cost of management and governance for technology in terms of full-time employees headcount's, legal, or other professional services or audit requirements (i.e., How much do the systems, processes, and technologies that process personal data cost?).

Model 2

Track time to deployment or proof of concept in a privacy engineering instance vs. traditional deployment. Start and track improvements in development, speed to deal closure, or other processes to attempt to measure organizational efficiencies.

Model 3

Work within the grain of cyber insurance. An enterprise will only be covered by cyber insurance where certain conditions are met to prove that the enterprise has taken at least reasonable steps to prevent loss. Create a checklist for coverage for various relevant scenarios based on the current level of cyber coverage or similar coverage within a relevant industry or size of enterprise for incidents such as hacker or other criminal external compromise, advanced persistent threat (APT) exploitation, negligent loss of media device, or physical encroachment. Generate the cost of repair or staffing to attain reasonable coverage in the event of a cyber incident.

Model 4

Look for qualitative or reputational examples rather than numerical values. For example, there are tools and techniques leveraging other individual's expressed curiosity, socially networked assertions, or trends according to big datasets or other analytics that can show relevance to the enterprise and value to individual customers.

Model 5

Leverage the known unknowns of brand valuation. Brand value determination is calculated using certain evidential or inferential techniques. Roughly stated, brand is measured as the difference between book value (adding all countable assets such as real estate and improvements, manufacturing assets, and the combination of financial assets relating to currency and investments) and market capitalization value. Where there is

a market and that market decides that a company is worth more than tangible assets, that differential is the collection of intangibles, potentials, and connective tissue that ties customers and employees to an enterprise and allows investors to decide an enterprise's potential.

PRIVACY IN THE ERA OF THE DATA ECONOMY

Chenxi Wang, Ph.D. Vice President of Market Insights at McAfee

We are living in the era of the data economy. The advent of consumer mobility and social media gave rise to a massive amount of readily available data to mine, aggregate, share, and analyze. IDC estimates that by 2020, there will be "40 zettabytes of information in the digital universe".[10] What's more, the composition of data products and applications can lead to brand new business models and previously impossible value propositions : consider Uber (the private, on-demand car service) in a world without Google maps.

Modern businesses now understand that access to data equals power and competitive advantages , and there is an increasingly large appetite to collect, store, and mind data. It is entirely possible that soon we will see a global market where data products and applications are routinely traded and exchanged. This trend has led to data obesity, heightened risk for data misuse, and an increasing concern for the threat to privacy.

Just like any other market, the data economy is governed by supply-and-demand and a value/pricing framework. Privacy regulations, however, typically seek to govern the supply and demand relationships, while completely ignoring the value framework. We argue that privacy is not attainable unless the value/pricing framework takes privacy impact into consideration. In other words, the value assessment of the data should not be solely based on their potential for creating valuable data products, but also based on their potential exposure to privacy risks.

Consider, for example, the case of a patron entering a bar. To gain admittance, today the patron needs to show her driver license, which discloses his date of birth, weight, height, and home address. Much of this information is beyond what the bar needs to know to permit entrance to the premises.

Consider again the same case when the patron approaches the bar, she is presented with three options: a) minimum disclosure to gain entrance (i.e., prove that she is over 21, the legal drinking age), b) disclose demographic information (i.e., age, gender) for a drink coupon, and c) consent for location tracking and ad serving for a much larger drink coupon.

[10]IDC's latest Digital Universe Report, released in December 2012, estimates that the amount of digital data produced will exceed 40 Zettabytes by 2020. This assumes all data is expected to double every two years.

If the patron chooses option A, her picture will be taken and sent to an information cloud for age verification. The answer that comes back from the cloud will be either a "yes (over 21)" or a "no (below 21)" , with no additional information such as date of birth. The picture is then deleted and the patron gains access to the premises.

If the patron chooses option B, the information cloud would disclose, along with age verification, demographic information such as age group, gender, etc. This information will be used in the bar operator's data mining and marketing efforts.

If the patron chooses option C, she would be asked to download an ad-serving app, which serves her relevant ads based on her location and activities.

Of course there could be other levels of information disclosure, but let's look at what just happened in the above scenario:

First, the customer has all the control: she can decide how much information to disclose.

Second, the marketers are not completely ignored here: they can get opt-in information, for a price.

The minimum disclosure is contextual: here the information disclosed is whether the user is above or below 21 years of age, but in other cases minimal disclosure can be about other data that make sense in the specific context of the activity. For example, location for local Yelp services may make sense in context.

There is a trusted intermediary—the info cloud in the example—that brokers the data exchange. The data broker does not have to be a singular party, but it needs to be a public entity trusted by the data owner.

To make this a reality, we need to establish a data value framework and a new model for the data supply chain. The data supply chain should include the designation of authoritative data suppliers, an access authorization model, authentication, data aggregation models, etc. The work done by UMA, for instance, is an example of an user-centric authorization model.[11]

The data value framework is arguably the most interesting, because it denotes how data will be assessed and traded, which are fundamental elements of an economy. One can consider a rudimentary value framework as follows: Pick your favorite data taxonomy, order the categories based on their exposure to privacy risks (if possible), and price them accordingly (the higher the risk, the higher the price tag). Afterwards, for each user-authorized data access, if the data required fall into minimal disclosure, they are supplied free of charge. Outside minimal disclosure, the data are

[11]UMA: User Managed Access (UMA) is an industry working group that is developing specifications that will allow an individual to control the authorization of data sharing and service access made between online services on the individual's behalf.

supplied with the attached price. For those data items that the user does not wish to grant third party access, the price tag can be set to infinite.

Clearly there are many options and intricacies to data value assessment beyond this basic framework. For example, how do you handle derived data, those that only exist based on previous data accesses? Similarly, the issue of what is considered minimal disclosure can be debatable.

However, we argue that without such a contextual data value model, either consumer privacy or the increasingly flourishing economy built on data sharing will be undermined. Businesses who truly understand the business impact of data and adopt this privacy-embedded data value framework will see consumers as willing participants in the data economy, where data exchanges are contextually relevant, properly priced, and in a manner that respects their privacy.

So, in many ways, the formula under a brand-based methodology could be that "brand" is the superset where intellectual property (IP) plus personal information (PI) are significant subsets of that market-driven asset. It is also illustrative that countries such as the United Kingdom, France, Australia, and New Zealand allow for intangibles to be included as part of an enterprise's balance sheet.

Brand values have been used to defend against a hostile takeover, as an investor relations tool, and, sometimes, as a performance indicator for the long-term investor. International standards that allow for intangible values may be leveraged and borrowed to assist in documenting PI value for the privacy engineer. For example, the International Accounting Standards Board (IAS 38), UK Accounting Standards (FRS 10 & 11), and US Accounting Standards Board (FASB 141 & 142 under Generally Accepted Accounting Principles) all may be used to determine or infer acquired goodwill. If the analogy from brand value to a subset of PI plus IP value is to be considered, it should be carefully noted and considered that the concept of "impairment"—roughly, the extent to which the stated value does not reach market value for a market-based enterprise—also impacts the PI value.

Here, the process and practice of privacy engineering becomes conceptually very interesting. Part of the controversial nature of valuing intangible assets is where those assets defy measurement. Compliance for data protection measures can be similarly difficult to achieve where enterprise governance professionals are unaware where data reside and how it is actually processed, and they do not have a means with which to measure processing over time. Where privacy engineering practices are followed, data is managed from its earliest analysis, design, and instantiation throughout its lifecycle. In such systems, active management and impairments based on market perception or active risk taking using data assets can be known and tested.[12]

[12]For example: www.nysscpa.org/cpajournal/2002/0202/features/202fp.22.htm

PRIVACY MATTERS BLOG SERIES: QUANTIFYING REPUTATIONAL RISK[13]

By Michelle Finneran Dennedy, published on Jan 06, 2012

There are many kinds of risk: operational, legal, and reputational risk. Most large enterprise IT teams are comfortable and proficient at measuring operational risk. It features in reports as minutes of downtime, incidents of endpoint reimages, number of patches installed, hours of overtime.

Legal risk isn't that hard to handle, either. IT can draw on peers, auditors, and legal staff for expertise.

However, reputational risk seems to be a far more unfriendly concept. I find technical people typically consider reputation a soft science, a squishy topic that can't be measured. As a result, IT can't set goals, gauge progress, or claim success based upon "reputation," and product creators cannot specify requirements for "reputation." Because it can't be managed like other metrics, IT staff and technical business units may ignore or downplay reputational risk's potential impact on the business—and their roles in protecting it.

IT is Missing a Gigantic Opportunity

I believe you *can* measure—or at least approximate—reputation, applying metrics to the same influences that affect your customers and your C-Suite executives: news headlines and stock prices. If you count the number of published, reputation-buffeting events each month—the headlines in the email news summaries you receive from SC Magazine, for example—you can see what the public is talking about, and that dialog will affect the rise and fall of organizational stock prices. Reputation and market sentiment are huge factors in market valuation, which is something your CMO and CFO are tracking. Although your interest may be in the technical security side of the business, you can take actions to measure, manage, or mitigate reputational risk.

Building a Reputational Heat Map

Well before the mortgage crisis was discussed in the public and mainstream press, it was anticipated in whispers at investment community conferences and insider blogs. Eventually, and much too late for most people and the economy, it was covered in USA Today and other mainstream papers on the doormats of hotel rooms coast to coast.

[13]This blog entry is reprinted in its entirety from McAfee's external web site: http://blogs.mcafee.com/business/security-connected/privacy-matters

Security issues that affect risk appear first in smaller, insider places, too. Then they migrate to the mainstream, to NPR, the Washington Post, Wired, and Vanity Fair. (Look at Stuxnet references on Wikipedia for a great example of this sequence.) With enough mainstream angst, trends start to register on the regulatory radar—with the European Community, the Federal Trade Commission, and others. We experienced this pattern with behavioral marketing. Privacy advocates raised objections in 2005, well before the FTC published its principles for behavioral marketing in December 2007. We are still seeing news and blog coverage on this topic today as companies experiment and push the envelope leveraging new technologies and relationships.

By the time a security topic attracts a reporter in the mainstream press, you had better have a strategy for that problem. You should be able to brief your boss with an assessment of your business's risk, including the risk to your reputation.

This assessment is possible, but you need to be selective. Just as you don't want to read every log entry from your IPS, you don't want to attempt to assess all topics everywhere on the Net. Instead, think about YOUR audience and what they read—or you wish they would read. Look at two tiers of publications: mainstream media and online influencers, including blogs and news feeds. Sign up for emailed daily updates if they are available from the 3–5 most relevant sources. Also, if there is an "insider" conference, you can look at the session titles and monitor news summaries for perspective on what the industry thinks is hot.

Next, think about what risks would affect your business and its reputation. The tech bloggers today might be talking about SQL injection, advertising dollars, identity theft, or phishing. What is newsworthy for your audience? Would a successful hack at a competitor raise questions about your security? Would regulation banning use of cookies affect your service offerings? If yes, use these ideas to set up RSS feeds.

That's your pre-work. You should revisit these decisions at least once a year, or when your business or the markets change significantly.

Now, the ongoing process. Your workflow is to:

Notice topics that relate to your risks.

Count the number of times these topics are mentioned in headlines or news stories. Depending on your work style (and the frequency of the publications you are tracking), you might either jot down mentions as you see them or save these mentions in a file for review monthly.

Create a spreadsheet: rows are the topics, columns are the dates. In each cell, note the number of headlines or significant mentions. If you think it's going to be important, start to capture dates and publications (use links if you can) so you can back up your ideas. (Store this info somewhere else, not in the mention count cell, or you won't be able to convert to a chart.)

> Once a month, use the spreadsheet's charting function to generate a "heat map," an assessment of which topics have generated the most energy in the news.
>
> If a relevant topic has generated significant coverage in insider publications, there's a good chance it will reach the mainstream press. If you think this might happen, summarize your findings in a concise note to your boss and your security team. Include an overview of what the issue is, what the coverage has been so far, what the impact would be on your business, and what efforts might be appropriate to mitigate these risks.

Voila.[14] You have quantified reputational risk.

Do this well, and you will be prepared if and when you need to discuss ideas with others. Instead of coming in with only technical data about a problem, you can talk with your colleagues in the context of the risk landscape. You look more strategic and more business-oriented. You are doing more, considering more, and recommending risk management efforts that are proportional to security. This position supports IT's increasing need to do internal selling to non-IT people in order to get the right projects funded.

At a minimum, this exercise will keep your knowledge of the risk landscape current, and you will be more fun at parties. You can talk to non-security people about ideas that they will recognize and explain risks in terms that they can understand. Perhaps you will detect the next "mortgage crisis" level event in time to help a few people avoid its devastation.

Building the Business Case

Measurements are only science projects until they are leveraged for positive progress. A privacy engineer's innovation can be lost without a market into which to sell the goods and services created with these methodologies or, similarly, it can be lost where internal enterprise measures are not sustained for continued improvement that results in better knowledge governance.

One approach is to treat privacy engineering products, services, and processes as *intrapreneurial* opportunities. An *intrapreneur* is an innovator within an enterprise who takes on the responsibilities for creating and "selling" new techniques or even new privacy business units. To become successful, intrapreneurial teams must connect with executive and operational teams to fit new things into existing environments effectively.

[14]Okay, so nothing is that easy, particularly in the world of data privacy and security, but hyperbole is a gimmick and the "voila" was a dead giveaway that I was trying to be dramatic for effect.

For example, when talking to the C-Suite:

- They hate details
- They don't know about detailed data privacy laws
- They hate details
- They have never seen a data valuation model, but they do like cost/benefit analysis where benefits are costed out realistically and the cost side looks real
- They hate details

PERVASIVE RISK MANAGEMENT APPROACH FOR EFFECTIVE PRIVACY ENGINEERING

By Vidya Phalke, Chief Technology Officer , MetricStream

A comprehensive and sustained risk management program is critical for an enterprise's long-term sustainability and predictability. Risk management needs to be comprehensive across all facets of operational, financial, legal, regulatory, reputational, data security, and intellectual property risks. In addition, it needs to permeate into an organization in a *pervasive* and deep fashion. The basic recipe for this pervasive treatment of governance, risk, and compliance is created by putting together models—both qualitative as well as quantitative—so that decision makers in an enterprise can create a deep understanding of their risks and then use that understanding effectively for planning and managing the short- as well as the long-term objectives at each level of the organization.

Although pervasive risk management is a broader topic, I will use this book's privacy focus to describe a mechanism by which a quantitative model can be orchestrated that will help management of risk that is based on how well privacy risk is understood and managed. This same mechanism can then be extrapolated for other areas of risks to arrive at a pervasive risk management architecture.

Whether it is government agencies or private organizations like banks, insurance companies, or health care providers, the need for incorporating privacy protection and managing privacy risk is not only a regulatory and legal obligation but it also has to be part of the risk management plan. The first step in tackling this risk is to create a comprehensive list of enterprise-wide assets and processes and map them to their privacy risk. This exercise typically is done in conjunction with IT and various functional units. If an enterprise already has a risk or compliance office, then that is usually a good place to start.

Second, a comprehensive assessment across all these assets and processes should be done along the dimension of privacy from a risk as well as a regulatory standpoint. If that has been done already, then that assessment can be leveraged.

The key here is to look for *privacy component* capabilities as described in this book in each of the assets and processes. In addition to looking for those components or their surrogates, a review of past audits, control or compliance testing, industry events or incidents, and other management evidence needs to be taken into account as well. Remember, this assessment needs to be wrapped into overall change management processes and frameworks.

Typically risks due to privacy issues will flow into both legal and regulatory risk as well as reputational risk, and assessment of the likelihood and impact has to be done based on qualitative and quantitative factors followed by evaluation of mitigating controls. As discussed in this book, the assessment need not be extremely precise but can start with an approximation. For example, measuring the privacy controls and usage of privacy components (or lack thereof) can lead to a score ranging from 0 to 5. These scores multiplied by the value of the asset or process they are tied to creates a weighted risk score. The process of assigning value to an asset or process in an enterprise is a well-defined science, so I will not spend time on that here; however, it suffices to state that it is tied to business criticality, footprint, and extent of being proprietary. For example, a database that contains PI that is accessible to an outsourced data analysis company will have a much higher footprint weight as compared to one that is accessible to a fixed known set of data analysts that are internal employees of that company. Once the comprehensive asset and process privacy risk assessments are computed, they need to be multiplied by the organizational weight of that business unit or functional division and then rolled into a score visible to the senior management.

Once this quantitative framework for risk management is put together, the next important aspect is to ensure that it is brought under the umbrella of enterprise change management; this is critical to ensure that as changes happen and new information is discovered, the impact of those changes is captured in the risk management framework. For example, in the above case of a database with PI, if a new application is being brought in that will be integrated within this database and will expose the data to a bigger set of users, then the risk parameters need to be reassessed and appropriate mitigation and controls need to be updated.

Figure 13-1 presents a pictorial summarization of this architecture and flow that should be applied to assets and processes to build a pervasive risk management framework and system.

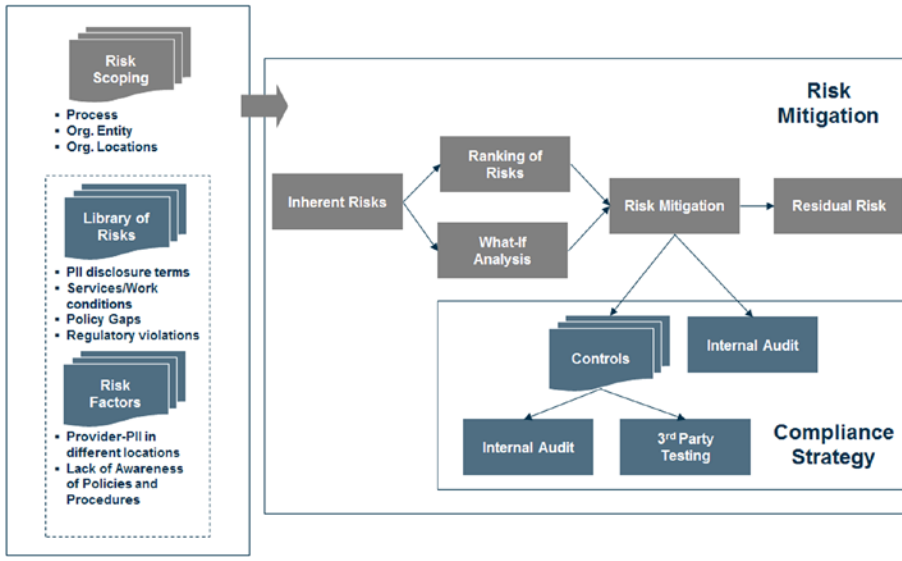

Figure 13-1. *Risk management with privacy use case*

Turning Talk into Action

Allies and other enterprise sponsors can help add to value models and create momentum. Privacy engineers *must* find allies such as the CFO, auditor, CMO, CTO, or any other leader willing to innovate with them and take on a bit of personal credibility risk. New things such as data valuation models can be perceived as unnecessary or not impactful or already managed by audit committees or compliance teams. Innovation in valuation models may require as many facts being marshaled from various measurement techniques as possible before a persuasive technique is selected for the enterprise.

Conclusion

The word "privacy" creates a marketing challenge. The paradox for creating data value models and systems can begin with this marketing issue. If enterprise stakeholders do not perceive or measure data risks and opportunities, they may well fall into a common trap. They may falsely assume that there is no need for privacy (after all, everyone says so). Another false assumption, if they do understand data about people or data derivatives have value, true stakeholders may feel that "someone else" owns or is accountable for the issue. Both false assumptions also suppose that data value is a thing or a static object as opposed to a flow, as is the case in capital- or currency-based value systems.