Smart Data Discovery and Transparency for Complying with CCPA, GDPR, and Other Privacy Regulations

David Stodder
TDWI Senior Director of Research, BI
March 19, 2019
Agenda

- Data privacy concerns and statutory response
- GDPR: Leading the way
- CCPA: Definitions and key provisions
- Data management considerations for complying with CCPA
- Recommendations
- Poll Question
- Audience Q&A
Rising Concerns About Data Privacy

• **Issue of trust**
  – Data breaches more common
  – Consumers becoming wary of being blind to what is being done with the data they generate
  – Fear of shared data being stolen, mangled, and/or inaccurate
  – Damage to reputations

• **Going into hiding**
  – Reports that consumers purposely falsify data to protect themselves and avoid marketing
  – Desire to be “forgotten”
Stepped Up Legal/Regulatory Pressure

- U.S. Federal Trade Commission: Rising number of enforcement actions regarding Internet privacy
- Massachusetts, Vermont, Colorado, and many other states stiffen consumer data protection laws
- Growing momentum for national privacy law
  - Reduce confusion of each state having its own law
2018: The Year of GDPR

• The European Union’s General Data Protection Regulation (GDPR)
  – “The most important change in data privacy regulation in 20 years”

• Ratified April 2016; into effect May 25, 2018
  – Reports indicate companies are still struggling to meet requirements

• Big influence on other countries and regions
  – Canada, Australia, Japan, China – and California
What GDPR Accomplished

- **Objective**: Protect rights of EU consumers
  - For transactions that occur within 28 EU member states, also exportation of personal data outside EU

- Provided set of regulations addressing how organizations capture, control, and process personal information

- Harmonized and strengthened data protection across EU; USA taking note
Redscan Freedom of Information Act (FOI) finding from analyzing 182 reports to U.K. Information Commissioner’s Office (ICO), as reported by www.helpnetsecurity.com (in the year before GDPR)

• Delays in reporting breaches to the ICO: Typically between 21 and 142 days (GDPR requires reporting within 72 hours!)
• Failure to include important information about the impact of the breach and the length of the recovery process (most did not know)
• Less than 25% of businesses would be GDPR-compliant as of April 2018
• Most companies lack skills, technology, or procedures to detect and handle breaches; identifying data breach “like finding a needle in a hay stack”
• Most data breaches happen on a weekend
Focusing the Mind: Stiff Penalties for GDPR Noncompliance

- Fines for not reporting personal data breaches within 72 hours (or for concealing)
- Penalties up to 4 percent of global annual revenues
- **Cost of compliance**: PwC survey finds most companies expect to spend between $1 million and $10 million to meet GDPR requirements (9% expect to spend more)
What It All Means: Organizations Cannot Be Passive

• Anticipating, discovering, reporting, and remedying data breaches must become part of an organization’s strategy
  – Increasingly, headed up by the CIO, chief data officer (CDO) or chief privacy officer (CPO)
• Organizational reputations at stake
• Attention on privacy will force changes in how organizations collect, process, and use data
GDPR: Assigning Accountability for Guarding Data

- **Data controller**: entity who determines purposes and means of processing personal data

- **Data processor**: entity who processes personal data on behalf of the controller
  - Including cloud providers

- Controllers and processors can be held liable for data breaches and noncompliance

- **Data protection officer**
  - GDPR requires controller and processor to designate a data protection officer (DPO) if they process or store EU citizen data
Now On to CCPA: The Clock is Ticking

- California Consumer Privacy Act (CCPA)
  - Enacted June 2018; becomes law January 1, 2020
  - Still being revised and amended (SB 561), including to impose stiffer penalties and protections
  - Protecting CA customers; affecting CA-based, for-profit companies and those with customer/consumers in CA

- Inspired by GDPR and based on a history of previous California laws

- “Fundamental to the right of privacy is the ability of individuals to control the use, including the sale of their personal information.”
CCPA: Size and Scope Are Big

- California: about 40 million residents
- Applies to for-profit companies with:
  - Revenues of $25 million or more
  - Hold data generated by 50,000 consumers, households, or devices
  - Derive 50% or more of annual revenues from selling consumers’ personal information
- 12-month “look back”: Need to report on records of PI collected dating from January 1, 2019
  - Restaurant chain example: collecting data collected via POS, customer use of Wi-Fi
  - Need to track what has changed; must monitor data sold or disclosed to third parties
What is “Personal Information”?

- **GDPR**: “Any info relating to an identified or identifiable natural person, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.” (www.gdpreu.org)

- **CCPA**: “Information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household.” (statute)

  - Many possible identifiers, including IP addresses, browsing history, inferences drawn from any of the identified information that creates a profile (www.blackpoint-it.com)
Important Provisions of CCPA

• **Right to know and access their PII:** Once a year and for free, companies have to tell CA citizens what categories of information they have collected on you, your devices, and your children

• **Sale or disclosure of data:** Requires businesses to tell CA citizens what categories of PI they are selling, and to whom
Important Provisions of CCPA (2)

• Right to say “no” to sale of PI: opt-out, vs. GDPR opt-in basis
  – Under 16 needs written opt-in
  – Businesses do not have to comply with request for deletion if there is a need to maintain the data in order to complete transaction, maintain adequate cybersecurity, repair service functionality errors, exercise free speech, or comply with CA Electronic Communications Privacy Act

• Equal service and price: If CA citizens opt out, cannot charge more or offer less
  – But can offer incentives for explicit opt-in
Penalties and Probable Revisions

• **Enforcement**: by CA Attorney General
  – *Probable revision*: Limits on resources make it impractical for AG to provide individualized guidance to companies; AG would be able to just give general guidance via publication
  – New bill would remove “right to cure” – the thirty-day remediation period before action can be taken. AG could enforce without prior warning

• Civil penalty for each violation can be up to $7,500
  – 10k customers = $75 million

• **Private right to action**
  – Revisions to CCPA expand permission to sue for violations and recover damages
PI Considerations for CCPA

• What PI is being collected and stored, and for what purpose?
  – Is it secure and its use well-governed?

• What categories of PI are being collected, shared, or sold?

• Status of communication with customers and consumers about opt-out privileges; with third parties

• Preparation for adherence, with expected revisions to statutes
First Question: Where Is Your PI Data?

- **Needed:** an inventory and complete view of customer/consumer data
  - Especially challenging when data spread across multiple platforms and cloud instances
  - Metadata useful to reveal PI “inferences” drawn from information gathered to create a profile about a consumer or customer

- **Data discovery:** Essential capabilities for locating the data and understanding data relationships
  - Technology to build awareness of inventory
  - Ad hoc, manual efforts cannot scale or keep up with requirements
Discovery for Determining Vulnerability

- Use of search to navigate data quickly
  - Concept search to capture relevant data relationships
  - Discover attributes and differences in how data has been recorded (often data quality issues)
- Discover data ownership and source origins

Cloud-native data storage, data lakes, and data warehouses

On-premises data warehouse and BI/OLAP data marts

Application databases (OLTP, Web, etc.); development and test environments
AI and Machine Learning (ML): Critical

• Technologies for data discovery and populating data catalogs need AI and machine learning
  – To address data volume and complexity
  – To work with streaming data

• ML for unbiased examination of data, indices, and more to generate view of data location and relationships
  – Critical where documentation is poor
  – Aids discovery across multiple systems
Data Discovery for Governance

- Governance rules and policies must address CCPA and other privacy and data protection statutes
- Data discovery for learning about and documenting data lineage
  - GDPR “right to an explanation” from analytics
- Governance and auditing over time
  - Ongoing data discovery to monitor PI data changes, including how it flows through business processes
  - Monitoring use in self-service BI and analytics
Governance & Data Discovery for Regulations

• Meeting compliance reporting and auditing requirements
  – GDPR and CCPA demand auditable processes; need visibility into data to monitor status

• Risk monitoring and reduction

• Handling consumer rights requests
  – Ensuring that right to say no is addressed in sales, marketing, and other processes
Recommendations

• Understand CCPA requirements
  – Headed by CDO, CPO, and/or governance committee
  – Align governance policies with statutes; look for overlaps with other regulations (e.g., GDPR, HIPAA, etc.)

• Take risk-based approach
  – Find and mitigate data vulnerabilities

• Establish how you will deal with requests

• Honor CCPA principles in business processes
  – Right to know, right to notification, to say “no” and opt out, nondiscrimination
Recommendations

- Improve data discovery for regulatory compliance
  - Evaluate AI/ML-driven technologies that can scale for data volume and complexity to locate and manage PI
  - Plan for ongoing discovery throughout data lifecycles

- Improve data cataloging
  - AI/ML is key here, too
  - Search, discover, and document data similarities and relationships

- ...and be prepared for change!
  - Regulations and statutes are always evolving
Thank You!

David Stodder
Senior Director of Research for Business Intelligence
TDWI (www.tdwi.org)
dstodder@tdwi.org
(415) 859-9933
Poll Question

• How far along is your organization on its journey toward compliance with data privacy regulations such as CCPA and GDPR?
  – We are fully prepared to comply
  – We have a plan and can locate most PI, but we need better technology, skills, and practices to complete our compliance
  – We can comply for some of our data but do not know where all the PI exists yet
  – We are just getting started and are hitting challenges
  – Don’t know or N/A
QUESTIONS?
CONTACT INFORMATION

If you have further questions or comments:

David Stodder, TDWI
dstodder@tdwi.org

Io-Tahoe, Webinar sponsor: Lola Bhadmus
lola.bhadmus@io-tahoe.com
Learn More in Chicago!

**TDWI Conference**
*Keynotes, Educational Classes, Networking, and More*
Chicago, Ill, April 28 – May 3, 2019
http://www.tdwi.org/chicago

**TDWI Strategy Summit**
*Case Studies, Expert Talks, and Leadership Panels*
Chicago, Ill, April 29-30, 2019
https://tdwi.org/events/strategy-summits/chicago