CCPA/CPRA: Implications for AI, Data Privacy, and Federated Learning.

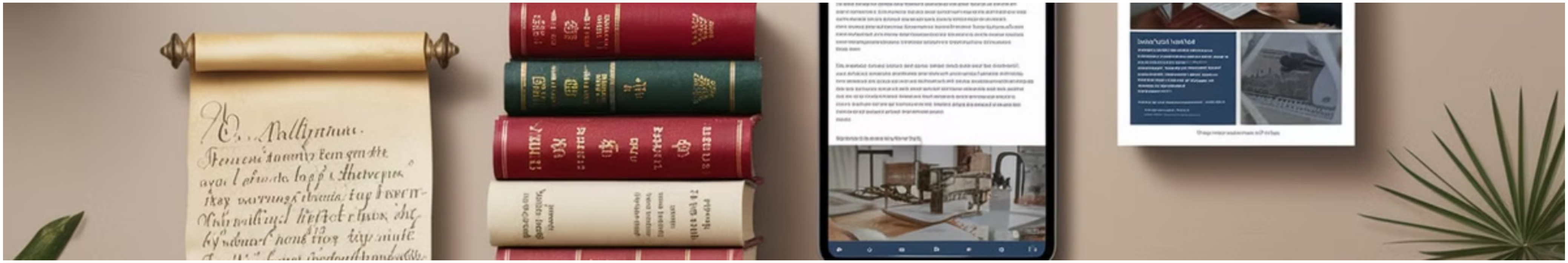
This presentation explores the transformative impact of the California Consumer Privacy Act (CCPA) and the California Privacy Rights Act (CPRA) on Artificial Intelligence (AI) and Federated Learning. We will examine key provisions, compliance strategies, and future trends.

Miguel (Mike) O. Villegas
CISO, AC|CISO, CTO, CISSP, CISA, CDPSE, C|EH, CSX|F, CSX|A, ISO27001 Lead Implementer
mike.villegas@isecureprivacy.com
213.453.6174

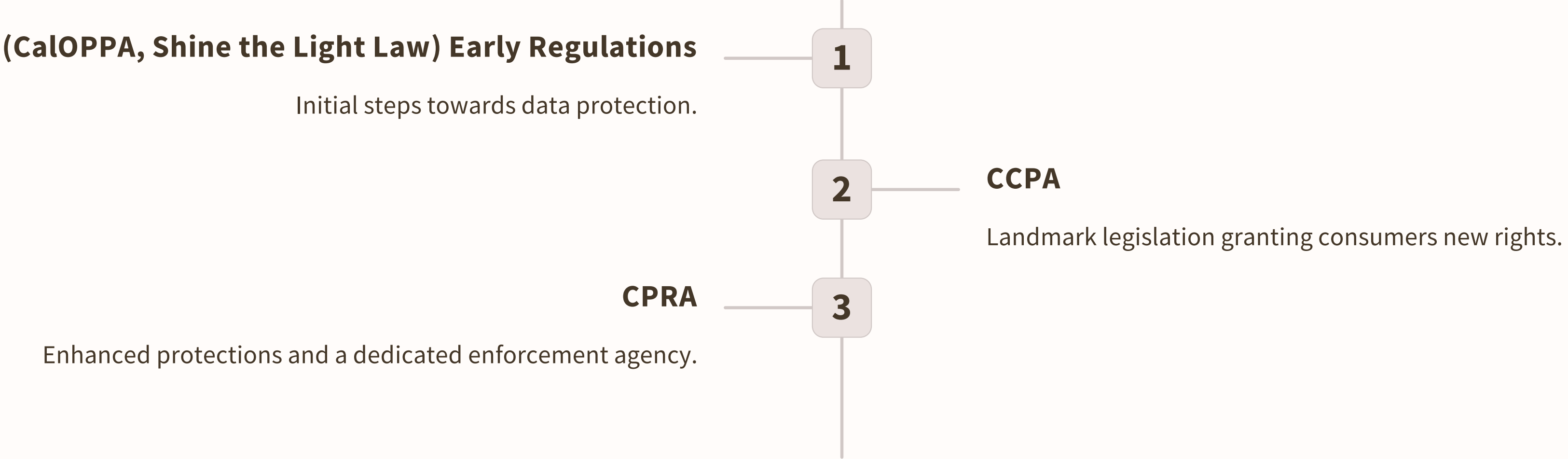
ABSTRACT

As artificial intelligence (AI) keeps advancing its impact across industries it becomes essential for IT professionals as well as web developers and cybersecurity experts to grasp how data privacy regulations relate to innovative AI methods. This lecture will analyze both the California Consumer Privacy Act (CCPA) and the California Privacy Rights Act (CPRA) with particular attention to how these laws affect AI technologies and data management methods and the developing concept of federated learning.

Participants will understand the fundamental concepts of the CCPA/CPRA including the rights of consumers and responsibilities of organizations handling data collection and processing activities. These regulations create substantial challenges for AI deployment because they require transparency in automated decision-making processes and proper use of consumer data.



Evolution of Data Privacy Laws in California



1 **(CalOPPA, Shine the Light Law) Early Regulations**

Initial steps towards data protection.

2 **CCPA**

Landmark legislation granting consumers new rights.

3 **CPRA**

Enhanced protections and a dedicated enforcement agency.

California has been a leader in data privacy, continuously evolving its laws to address emerging challenges and protect consumer rights in the digital age.

Key Objectives of CCPA and CPRA

1 **Enhance Consumer Rights**

Ensuring consumers understand how their data is collected and used.

2 **Strengthen Business Responsibilities**

Transparency, Data Minimization, Security Obligations, Vendor TP Controls

3 **Establish Dedicated Privacy Enforcement Agency**

Holding businesses responsible for data protection practices.

4 **Expand Business Coverage**

CPRA lowered threshold for covered businesses for 100,000 consumers

These acts aim to empower individuals with greater control over their personal information and increase transparency in data handling practices.

Scope

1. **Businesses Subject to CCPA/CPRA**

   CCPA/CPRA applies to **for-profit businesses** that meet **any** of the following criteria:

✅ **Revenue Threshold**

Businesses with **annual gross revenues of $25 million or more**.

✅ **Data Processing Volume**

Businesses that **buy, sell, or share the personal data of 100,000 or more California residents, households, or devices per year** (CPRA increased this from 50,000 under CCPA).

✅ **Revenue from Data**

Businesses that **derive 50% or more of their annual revenue from selling or sharing consumers' personal information**.

✅ **Geographic Scope**

Even if a company is **not physically located in California**, it **must comply** if it processes data of California residents and meets the above criteria.

2. **Service Providers & Contractors**

✅ **Service Providers**: Businesses that process data on behalf of another company, such as cloud providers, advertising agencies, or outsourced IT services.

✅ **Contractors**: Companies that receive personal data but are restricted in their use of it (e.g., payroll processors, cybersecurity firms).

🔷 **Why It Matters?**

•Service providers and contractors **must follow contractual obligations** ensuring they **only use personal data for specified purposes** and comply with privacy regulations.

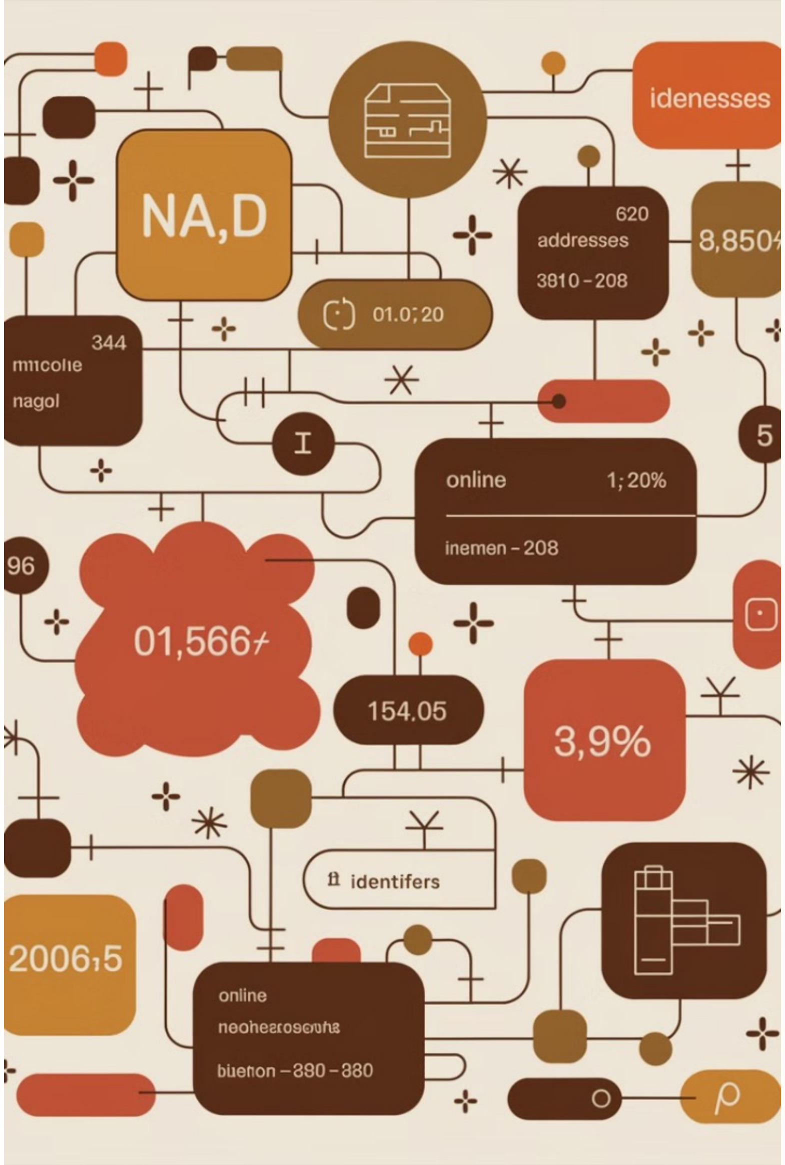
3. **Consumers & Employees**

✅ **California Residents**: The law applies to **any resident of California**, regardless of where the business is located.

✅ **Employees & Job Applicants** (CPRA Expansion):

•CPRA **extended consumer rights to employees**, job applicants, and independent contractors.

•Businesses must provide **data access and correction rights** for employee data starting in **2023**.



Definition of Personal Information Under CCPA/CPRA

Broad Definition

Information that identifies, relates to, describes, or is capable of being associated with a consumer.

Examples

Name, address, IP address, email address, browsing history.

The definition of personal information is comprehensive, encompassing a wide array of data points that can be linked to an individual.

Consumer Rights Under CCPA/CPRA

Right to Know

Access information collected about them.

Right to Delete

Request deletion of personal data.

Right to Opt-Out

Prevent sale of their data.

These rights empower consumers, giving them control over their personal information and how it is used by businesses.

Business Obligations and Compliance Requirements

1 **Data Inventory**

Identify and categorize personal data.

2 **Privacy Policy**

Provide clear and transparent information.

3 **Training**

Educate employees on compliance.

Businesses must implement comprehensive measures to ensure they are meeting the requirements of CCPA/CPRA and protecting consumer data.

Special Categories of Sensitive Data

Definition

Data that requires extra protection due to its sensitive nature.

Examples

Financial information, health data, precise geolocation.

CCPA/CPRA places stricter requirements on the handling of sensitive personal information, ensuring heightened privacy safeguards.

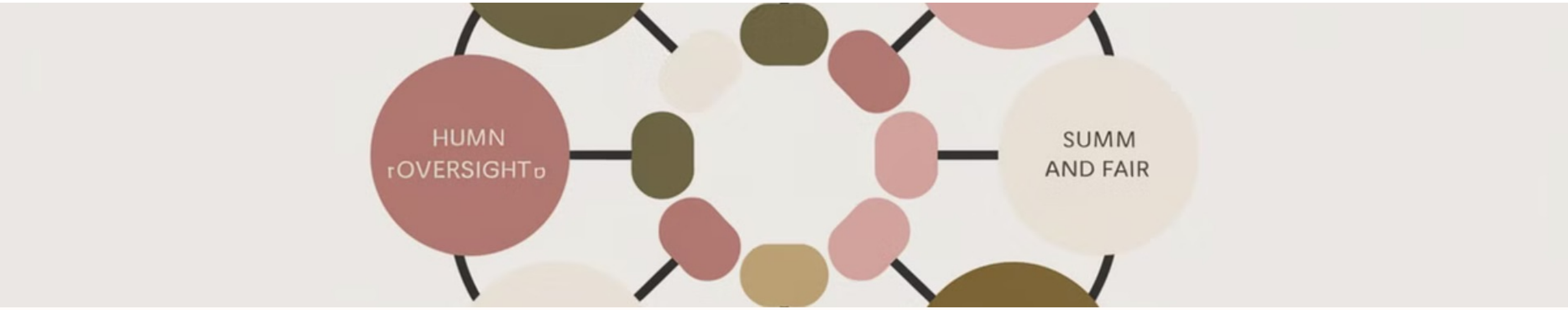
Impact on AI Development and Implementation

Data Collection and Usage Limitations

Consumer Rights and Automated Decision-Making

Key Considerations for IT Personnel, Web Developers, and Cybersecurity Professionals

AI systems must be developed and implemented in a manner that respects consumer rights and complies with privacy regulations.



Data Collection and Usage Limitations

1. **Data Requirements for AI Training**
2. **Limitations on Data Sharing**
3. **De-Identification Techniques**

Businesses must be transparent about their use of automated decision-making technologies and provide consumers with options.

Data Minimization Principles

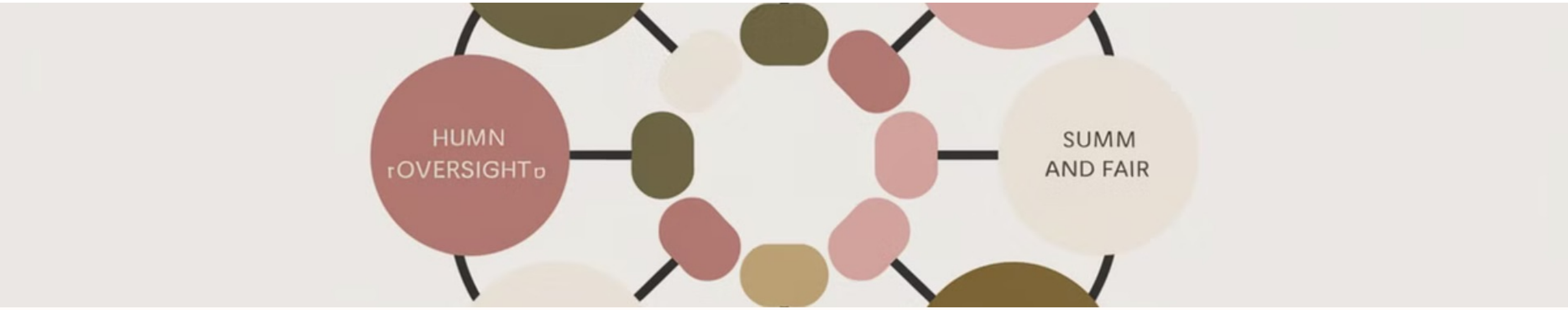
Collect Only Necessary Data

Limit collection to what is relevant and proportionate.

Purpose Limitation

Use data only for specified purposes.

Adhering to data minimization principles reduces privacy risks and promotes responsible data handling.



Consumer Rights & Automated Decision Making

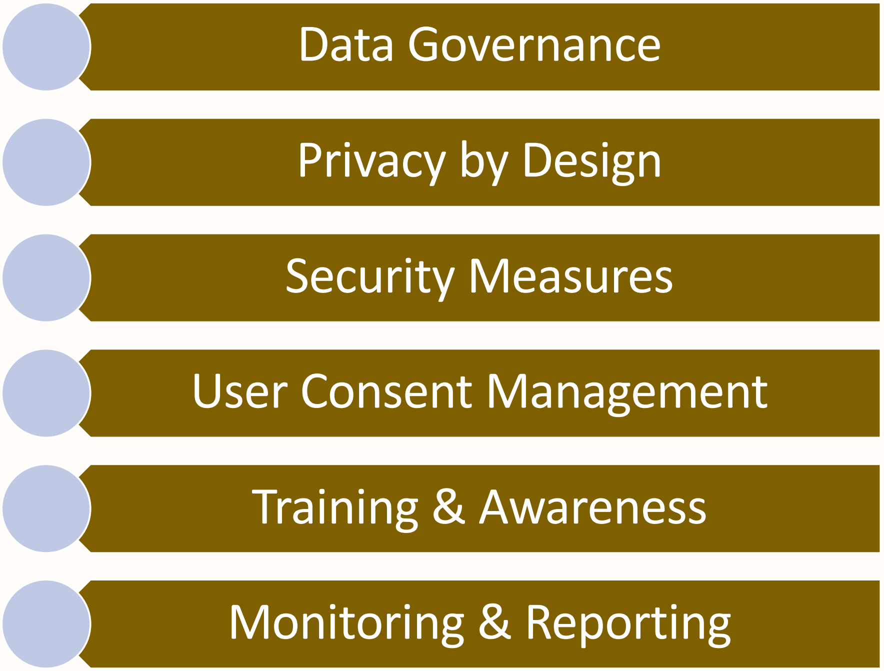
1 **Transparency in Algorithms**

2 **Impact Assessments**

3 **Right to Explanation**

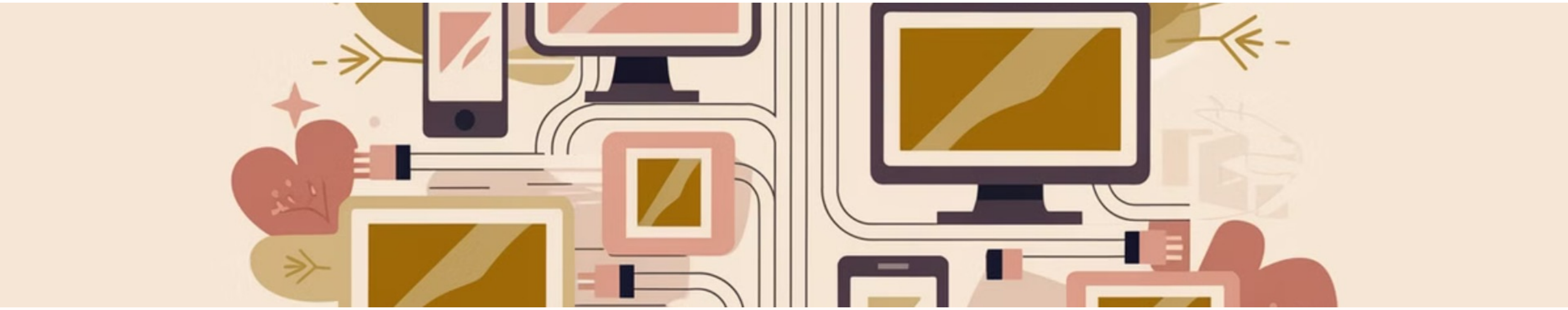
Businesses must be transparent about their use of automated decision-making technologies and provide consumers with options.

Key Considerations for IT Personnel, Web Developers, and Cybersecurity Professionals



- Data Governance
- Privacy by Design
- Security Measures
- User Consent Management
- Training & Awareness
- Monitoring & Reporting

CCPA/CPRA imposes significant responsibilities on organizations using AI, necessitating a proactive approach to data privacy, security, and compliance.



Federated Learning

1. Decentralized Data
2. Local Training
3. Model Aggregation
4. Privacy Preservation
5. Reduced Bandwidth Usage

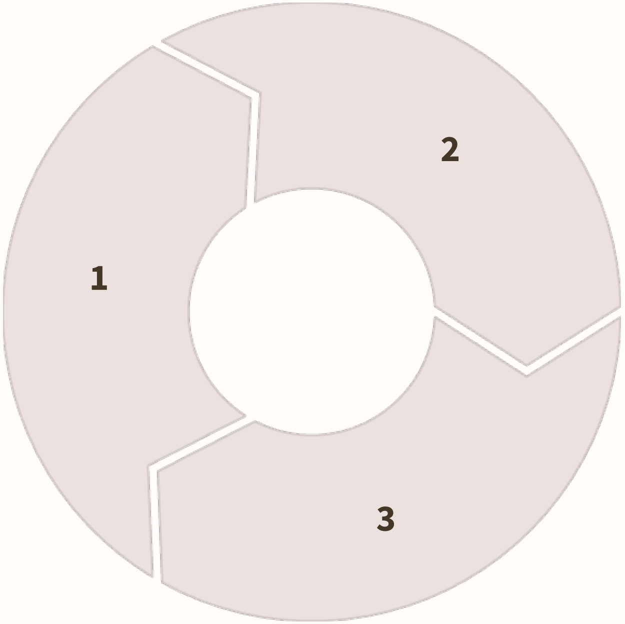
Applications of Federated Learning

Healthcare	In healthcare, federated learning can be used to train models on patient data across different hospitals without sharing sensitive health records. This allows for improved predictive models while maintaining patient confidentiality.
Finance	Financial institutions can use federated learning to develop risk assessment models using transaction data from various branches without exposing individual customer data, thus enhancing fraud detection while adhering to privacy regulations.
Smart Devices	Federated learning is commonly used in mobile applications, where models are trained on user data from smartphones. For instance, predictive text input features in smartphones can improve based on users' typing patterns without sending individual text data to the cloud.
Autonomous Vehicles	In the automotive industry, federated learning allows vehicles to share insights without sharing raw sensor data, enabling collaborative learning for better navigation and safety systems.

Challenges and Considerations

Data Heterogeneity	Data across devices can vary significantly in quality and quantity. This heterogeneity can affect model performance, as some devices may have more representative data than others.
Communication Effectiveness	The communication overhead between devices and the central server can be a bottleneck. Optimizing the frequency and size of updates is essential to improve efficiency.
Security Concerns	While federated learning enhances privacy, it is not immune to attacks. Techniques such as differential privacy can be implemented to add noise to the model updates, further protecting individual data points from being inferred.
Model Convergence	Achieving convergence in federated learning can be more complex than in traditional methods due to the decentralized nature of training. Researchers are actively exploring algorithms to enhance convergence rates.

Data Processing Agreements and Vendor Management



Due Diligence

Assess vendor privacy practices.

Contractual Terms

Include data protection clauses.

Monitoring

Regularly monitor vendor compliance.

Businesses must carefully manage their vendors and ensure they adhere to CCPA/CPRA requirements through robust data processing agreements.

Risk Assessment Requirements

1 **Identify Risks**

Assess potential privacy risks.

2 **Implement Safeguards**

Mitigate identified risks.

3 **Regular Review**

Update assessments periodically.

Regular risk assessments help identify and address potential privacy vulnerabilities, ensuring ongoing compliance with regulations.

Privacy Impact Assessments

PIA helps organizations **identify, assess, and mitigate** privacy risks associated with the collection, use, and processing of personal data in AI systems.

It ensures compliance with privacy regulations like **CCPA/CPRA, GDPR, and other data protection laws**, while also fostering transparency and consumer trust.

AI models rely heavily on **large-scale data collection**, making them prone to **privacy risks, bias, and security vulnerabilities**. A well-conducted PIA helps:

- Identify potential **privacy risks** before deploying AI systems.
- Ensure **compliance** with data protection laws.
- Promote **transparency and trust** in AI decision-making.
- Prevent **legal and reputational risks** due to non-compliance.
- Enhance **fairness and accountability** in AI-driven decisions.

Privacy Impact Assessment (PIA)

Personal Information

Categories

- Employee Data
- Customer Data
- Patient Data
- Suppliers/Vendors
- Company Information

- Identifiers such as a real name, alias, postal address, unique personal identifier, online identifier, Internet Protocol address, email address, account name, social security number, driver's license number, passport number, or other similar identifiers.
- Characteristics of protected classifications under California or federal law.
- Biometric information.
- Professional or employment-related information.

- Commercial information, including records of personal property, products or services purchased, obtained, or considered, or other purchasing or consuming histories or tendencies.
- Internet or other electronic network activity information, including, but not limited to, browsing history, search history, and information regarding a consumer's interaction with an internet website application, or advertisement.

- Geolocation data.
- Audio, electronic, visual, thermal, olfactory, or similar information.
- Education information, defined as information that is not publicly available
- Inferences drawn from any of the information identified to create a profile about a consumer reflecting the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes.
- Sensitive personal information."

iSecurePrivacy LLC

OWASP

Privacy Impact Assessment (PIA)

Data Inventory

- Data – describe the data
- Source – where does it come from or generated
- Location/Where Stored
- Were has data been sent?
- Technology Used (Is the data located on:
  - (1) database (e.g., MS SQL, Oracle, MySQL, Postgres, etc)
  - (2) server (e.g., Windows, Linux, HP/UX)
  - (3) file share (Sharepoint, Windows, Google Docs, Box,etc),
  - (4) mobile device (e.g., laptops, smartphones, tablets),
  - (5) applications (e.g., Excel, MS Access),
  - (6) cloud services (e.g., AWS, Azure, Google Apps, Office 365)
- Data Purpose
- Consent of Data Subjects?
- Retention Requirements
- Parties with Access to Data
- Security Measures to Protect Data
- Destruction – how is it destroyed?
- Contact information of the Data Controller, Data Processor, and Data Protection Officer
- Classification / Sensitivity
- Business Criticality
- Comments

Consent Management and Dark Patterns

Informed Consent

Informed consent refers to the process by which users explicitly agree to the collection, processing, or sharing of their personal data **with full knowledge of the implications**. It is a **fundamental principle** in privacy laws like **CCPA, CPRA, and GDPR**, ensuring that individuals retain control over their data.

Avoid Dark Patterns

Dark patterns are deceptive design techniques used in user interfaces that **manipulate users into making unintended choices**—often to **collect more personal data, make opt-outs difficult, or pressure users into purchases**. These violate **privacy regulations** like **CCPA/CPRA**, which prohibit misleading consent practices.

Common Types of Dark Patterns

1. **Forced Opt-In (No Choice Given)**
- Users must **agree to data collection** to use a service, with no option to **decline**.
- Example: A social media platform **requiring** phone number verification without an alternative.

2. **Hidden or Buried Privacy Settings**
- Making it **hard to find** privacy controls or settings.
- Example: Opt-out options buried in **multiple pages of menus**.

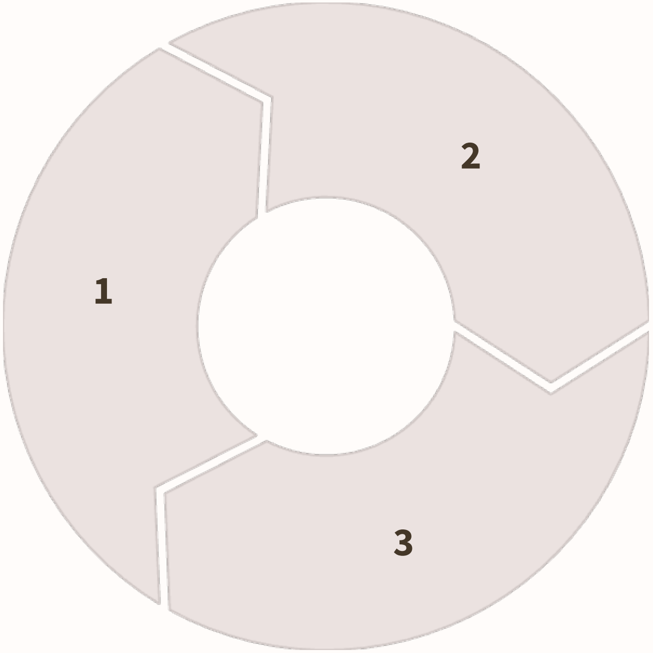
3. **Pre-Checked Boxes (Assumed Consent)**
- Users are automatically opted into data collection unless they **manually uncheck a box**.
- Example: A form where the checkbox for "Receive promotional emails" is pre-selected.

4. **Misdirection & Confusing Wording**
- Using misleading language to **trick users into agreeing** to data collection.
- Example: A **double-negative** opt-out: “Uncheck this box if you do not want to opt out.”

5. **Guilt or Fear-Based Messaging**
- Pressuring users into **giving consent** by making them feel guilty.
- Example: A pop-up that says:
  - *“By declining, you’re missing out on an amazing experience!”*
  - *“Are you sure? Your account may not work properly!”*

6. **Friction for Opt-Out, Easy for Opt-In**
- Making it **hard to cancel a subscription or decline tracking**, but easy to accept it.
- Example:
  - Opting **into** a subscription with **one click**.
  - **Canceling** requires **calling customer service** or **multiple verification steps**.

Enforcement Mechanisms and Penalties



CPPA

California Privacy Protection Agency enforces.

Investigations

Agency investigates complaints.

Penalties

Significant fines for non-compliance.

The CPPA has the authority to investigate and enforce CCPA/CPRA, with significant penalties for non-compliance.

Civil Penalties Under CCPA/CPRA

Violation	Potential Fine	Notes
Failure to provide opt-out options for data sales	$2,500 - $7,500 per violation	Applies to **selling/sharing consumer data**
Denying consumer data access or deletion requests	$2,500 - $7,500 per violation	**Consumers must have clear rights to access and delete their data**
Using deceptive **dark patterns** to obtain consent	$7,500 per violation	**Misleading opt-in tactics are strictly prohibited**
Unauthorized use or sale of **children’s data (under 16)**	$7,500 per violation	**Higher penalties for mishandling minors' data**
Data breaches due to **lack of security measures**	$100 - $750 per affected user OR actual damages	Consumers can file lawsuits **individually or collectively**

Best Practices for AI Compliance

1

Privacy by Design

Incorporate privacy from the outset.

2

Transparency

Be open about data practices.

3

Accountability

Take responsibility for data protection.

Implementing these best practices helps organizations build trust with consumers and comply with CCPA/CPRA requirements.

Implementation Timeline and Key Dates

1 **CCPA Effective**

January 1, 2020

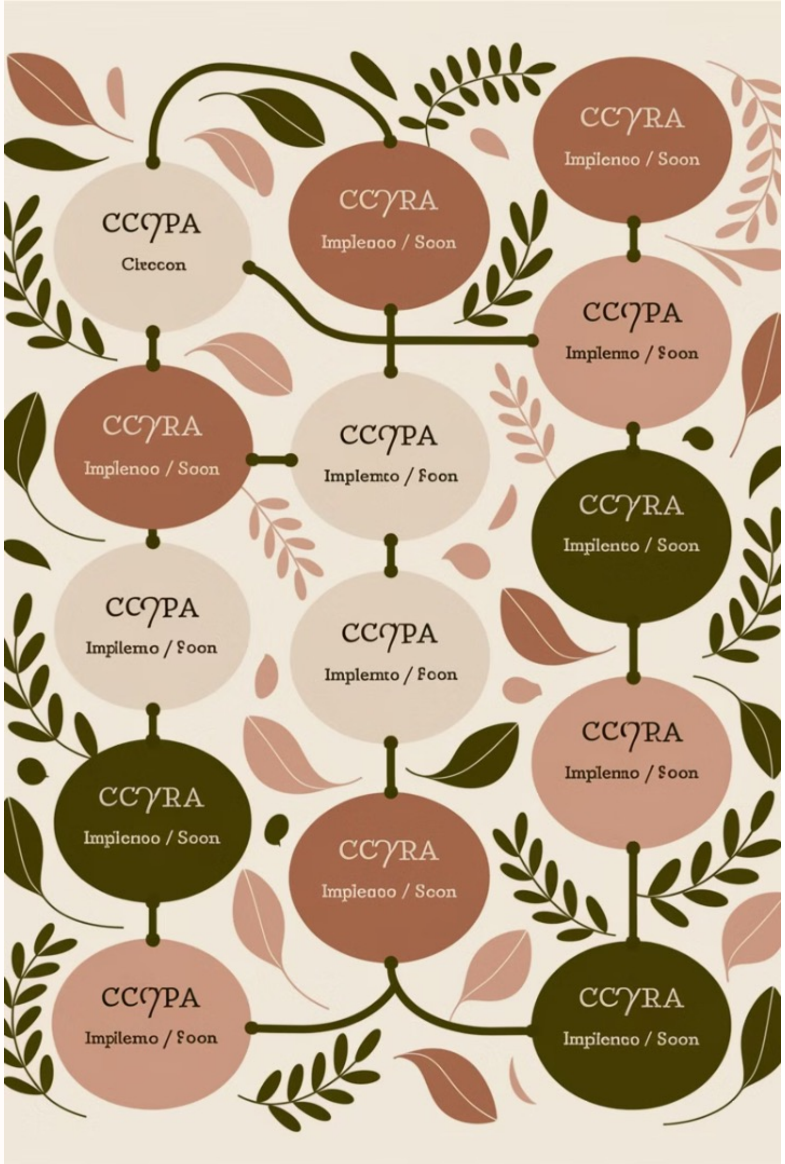
2 **CPRA Passed**

November 3, 2020

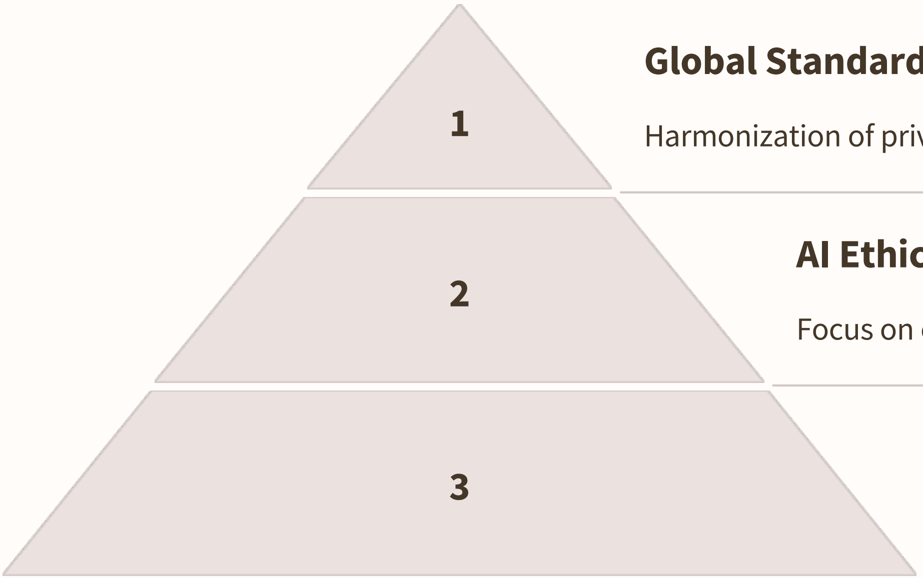
3 **CPRA Effective**

January 1, 2023

Staying informed about key dates and deadlines is crucial for ensuring timely compliance with CCPA/CPRA.



Future of AI Privacy Regulations



Global Standards

Harmonization of privacy laws.

AI Ethics

Focus on ethical AI development.

Technological Advancements

New privacy-enhancing technologies.

The future of AI privacy regulations will likely involve greater harmonization, a stronger focus on ethics, and the development of new technologies.

Action Items and Next Steps

1

Assess Compliance

Evaluate current practices.

2

Update Policies

Revise privacy policies.

3

Train Employees

Educate workforce.

Take these action items to ensure your organization is prepared for the evolving landscape of AI privacy regulations and can protect consumer data.

Miguel (Mike) O. Villegas

Mr. Villegas is the Founder and President of iSecurePrivacy LLC, a technology consulting firm focused on cybersecurity and privacy of critical risk information. Mr. Villegas is currently CISO for Tristar Insurance Group, the largest privately-owned Third-Party Administrator (TPA) serving the entire USA. He is also CTO and CISO for Xahive, a startup secure email solution and security awareness provider. He was previously a Senior Vice President for K3DES, a technology consulting firm focused on the security of electronic payments systems (PCI).

Mr. Villegas is a Certified Information Systems Auditor (CISA), Certified Information Systems Security Professional (CISSP), AC|CISO, CSX|F, CSX|A, Certified Ethical Hacker (CEH), and ISO/IEC 27001 Lead Implementer. He was the 2010-2012 President of the ISACA Los Angeles Chapter and the 2005-2006 President of the ISACA San Francisco Chapter. He was Co-Chair of the SF ISACA Fall Conference from 2002 through 2008.

He is currently the Certification Chair for the ISACA Los Angeles Chapter, a member of the LA Spring Conference Committee, and the COBIT Technical Review Committee for LA ISACA. He was also a Board Member for the ISSA Los Angeles, CA Chapter, and a member of ISSA, ISC(2), and OWASP. Mr. Villegas was also a contributing writer for SearchSecurity—TechTarget, with over 150 articles written.