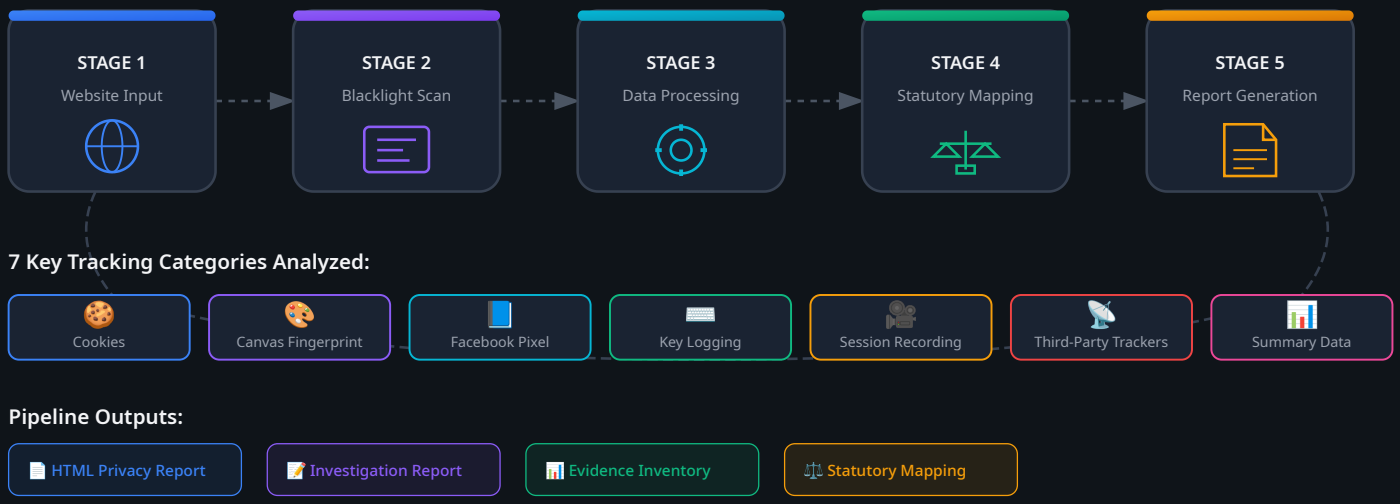
Privacy Analysis Pipeline

A technical assessment system for documenting website tracking practices and mapping detected data collection to California statutory definitions under CIPA, CDAFA, and Penal Codes § 631 & § 638.51



Pipeline Stages Explained

1 Website Input & Configuration

The pipeline begins with target website identification. URLs can be processed individually or in batch mode via CSV files. Configuration includes defining scan parameters, consent testing scenarios, and evidence collection requirements.

URL Input | CSV Batch Processing | runone.sh

2 Blacklight Privacy Scan

Automated crawling using The Markup's Blacklight scanner to detect tracking technologies. Captures cookies, canvas fingerprinting, session recording scripts, Facebook pixels, key loggers, and third-party tracker requests.

Blacklight Scanner | Puppeteer | Network Analysis

3 Data Processing & Parsing

Python scripts process raw Blacklight CSV outputs with comprehensive analysis. All cookies are analyzed (not just samples) with automatic delays and retry logic for API rate limiting. Data is normalized and categorized for legal framework mapping.

Python Scripts | CSV Processing | Data Normalization

4 Statutory Framework Mapping

Detected tracking practices are categorized according to California statutory definitions. Data collection is mapped to TRAP AND TRACE (§ 638.51) addressing information categories and WIRETAP (§ 631) content interception categories. This technical mapping provides documentation for subsequent legal review by counsel.

Evidence Categorization | Statutory Definitions | § 631 / § 638.51

5 Report Generation

Professional HTML reports are generated with legal-quality formatting. Reports include complete evidence inventories, specific cookie counts, tracker instances, and clear separation between technical findings and legal compliance evaluation.

HTML Reports | Markdown Docs | Evidence Package

Statutory Framework Mapping

§ 638.51 TRAP AND TRACE - Addressing Information

Data collection mapped to statutory definitions of "addressing information":

→ IP address collection and transmission

→ Geolocation data (including ZIP codes)

→ Unique device identifiers

→ Browser and device fingerprinting

→ Age and demographic information

→ User agent and system information

§ 631 WIRETAP - Content Interception

Data collection mapped to statutory definitions of intercepted "contents":

→ Page view and content tracking

→ Form interactions and submissions

→ Search queries and clicks

→ Session recording and replay

→ Key logging and input capture

→ E-commerce tracking (cart, checkout, purchase)

Tracking Technologies Detected

🍪

Cookies

First-party, third-party, session, persistent, HTTP, JavaScript

🎨

Canvas Fingerprinting

Graphics rendering exploitation for unique device identification

📘

Facebook Pixel

Standard and custom events for cross-site tracking

⌨️

Key Logging

🎥

Session Recording

📡

Third-Party Trackers

Keyboard input monitoring and capture scripts

FullStory, Hotjar, LogRocket and similar tools

Advertising networks, analytics platforms, data brokers

Final Deliverables

HTML Privacy Report

Professional, legal-quality formatted report with comprehensive findings and evidence citations

Investigation Report

Markdown documentation with detailed technical assessment and methodology notes

Evidence Inventory

Complete raw data dumps including all cookies, trackers, and fingerprinting instances

Statutory Mapping

Categorized findings mapped to TRAP AND TRACE and WIRETAP statutory definitions for legal counsel review

---