Analyse von Logdateien

- Webserver generieren Logdateien, die eigentlich zur Fehlerbehebung gedacht sind, woraus man aber unter anderem die grobe Nutzeraktivität entnehmen kann
- aufgrund von zusätzlich gesendeten Informationen kann man auch der Anonymisierung von IP-Adressen entgegenwirken (z.B. durch VPNs/Proxies)

Pixel/Beacons

- eines der ersten Verfahren
- sehr kleines Bild, das beim Aufrufen einer Webseite geladen wird
- gängige Methode für Lesebestätigung von E-Mails

Tag/URL-Tracking

- in der URL befinden sich bestimmte Parameter, die direkt vom Webserver oder lokal gesetzt werden
- diese Parameter werden bei allen Anfragen mitgeschickt und können so für Tracking verwendet werden

Cookies

= kleine Mengen an Daten, die der Browser bei Anfragen mit an den Server sendet

<u>Verwendung</u>

- Sitzungsverwaltung (Login, Einkaufskörbe, ...)
- Personalisierung (lokale Einstellungen wie Designs)
- Tracking (Verfolgung und Analyse von Nutzerverhalten)

Fingerprinting

- “digitaler Fingerabdruck”, viel invasiver als Cookies
- eindeutige Identifizierung geschieht durch Dinge wie Gerätehardware, -software, Browsererweiterungen und Einstellungen
    - z.B. Betriebssystem, Architektur, installierte Schriftarten, Bildschirmart und -größe, Browser, Grafikkarte & Treiber, uvm.
    - Canvas/WebGL Hash
        - Nutzung von HTML <canvas> Elementen bzw. WebGL
        - genaues Ergebnis hängt von extrem vielen, kleinen und unkontrollierbaren Faktoren ab
          -> nahezu eindeutig und damit sehr verlässlich
- im Gegensatz zu Cookies fest an Gerät gebunden, kann also nicht einfach gelöscht werden

Federated Learning of Cohorts (FLoC)

= vorgeschlagener Lösungsansatz von Google

- Tracking geschieht nicht extern auf Webservern mit Fingerprinting & Cookies, sondern direkt im eigenen Browser
- durch förderales Lernen (Art des maschinellen Lernens, die auf mehreren Geräten stattfindet) werden Kohorten (Interessensgruppen) lokal ermittelt und auf Anfrage mit Diensten geteilt

<u>Vorteile</u>

- erhöhte Sicherheit (Verschlüsselung)
- Nutzungsverhalten und Verlauf an sich bleiben privat und im Browser
    - ermöglicht trotzdem zielgerichtete Werbung
- alle Webseiten, die sich für FLoC entscheiden, haben Zugriff auf die gleichen Daten (sie sind somit nicht exklusiv)

<u>Nachteile</u>

- ohne explizite Zustimmung; man kann nicht entscheiden, welche Seiten auf die Daten zugreifen dürfen
- vereinfacht Fingerprinting enorm, da man statt dem ganzen Internet nur noch kleinere Interessensgruppen auseinanderhalten muss