¡Galletas!

La receta detrás de las cookies de Internet

Sistemas de la Información 2024/2025
Mario Clavero Artal 875371
Andrea Ferradas Agudo 873635
Laura Hernández Sierra 872203
Marta Ibáñez Alloza 869896
Jorge Hernández Aznar 872838
Paula Blasco Díaz 874320

Índice

1. Introducción

Hoy en día todo es conocedor del uso de las cookies, son un componente esencial para la personalización de los sistemas de información. Plataformas como Netflix las usan para las preferencias de visualización, recomendaciones según el historial personal previo o para mantener una sesión activa. Asimismo en sitios como Amazon son una herramienta fundamental para el carrito de la compra. No obstante, pese a este uso diario no todo el mundo es consciente de su origen o de su funcionamiento más técnico. Por este motivo hemos decido realizar el trabajo práctico de la asignatura sobre este tema, basándonos en una entrada pública en el blog de la bonilista.

Origen de las Cookies

En los años 90, la web era un entorno completamente diferente al actual. El contenido presentado mediante páginas HTML era únicamente estático y la posibilidad de interacciones se limitaba a los enlaces a otras. Por otro lado, existían otras alternativas como Gopher, un sistema jerárquico de archivos, o redes BBS, que ofrecían servicios restringidos a comunidades específicas. También destacaban redes cerradas como CompuServe, que limitaban el acceso a sus contenidos y funciones. Este funcionamiento inicial diseñado por Tim Berners-Lee tenía la idea de intercambio de información unidireccional.

En 1994 se fundó Netspace una empresa de software estadounidense que tuvo un gran impacto en el crecimiento de la web, aportando el inicio de los formularios y uno de los primeros buscadores de páginas web. Entre sus empleados se encontraba Lou Montulli, un joven desarrollador con experiencia en sistemas de información. La empresa recibió el encargo de realizar una aplicación de e-commerce por parte de la compañía de telecomunicaciones MCI. Para ello uno de los requisitos fue que el estado de las transacciones de compra no completadas no se almacenará en el servidor sino en el navegador de cada usuario.

Para ello a Montulli se le ocurrió la idea de implementar un mecanismo que permitiera a los navegadores conservar información entre distintas peticiones. Inspirado por el concepto de *magic cookies*, tokens utilizados en sistemas operativos para identificar programas o usuarios, Montulli desarrolló una solución basada en pequeños fragmentos de datos denominados cookies. Estas se almacenaban como un simple archivo de texto en el navegador del usuario y se enviaban automáticamente como metadatos junto con las solicitudes futuras al servidor correspondiente.

De esta manera la experiencia del usuario dentro de la web tenía una interacción más dinámica. Pasado un tiempo, el 13 de octubre de 1994, Netscape lanzó su versión 0.9beta de su navegador, este fue el primero navegador que usaba cookies y su principal objetivo fue comprobar si los visitantes ya habían visitado previamente el propio buscador. Desde ese momento las cookies comenzaron a consolidarse como una herramienta clave en la evolución de la web.

2. Funcionamiento técnico de las cookies

Las cookies son pequeños fragmentos de información que las páginas web almacenan en el navegador del usuario. Sirven para mantener el estado entre peticiones sucesivas de una misma sesión, permitiendo a los sitios recordar información sobre el usuario, como sus preferencias o la autenticación de su sesión.

Las cookies funcionan mediante un sistema de pares clave-valor, en el que la clave es el nombre de la cookie y el valor es la información asociada. Por ejemplo, sessionID=728. Estas cookies se envían entre el navegador y el servidor en los encabezados HTTP de las peticiones y respuestas. Para establecer una cookie, el servidor envía un encabezado Set-Cookie en la respuesta HTTP. Posteriormente, el navegador incluirá esa cookie en las solicitudes HTTP futuras mediante el encabezado Cookie.

Por ejemplo, en el contexto de una tienda online, cuando un usuario realiza su primera visita, el servidor podría responder con un encabezado Set-Cookie que contiene un identificador único de sesión, como un sessionID. Esta información se guarda en el navegador como una cookie.

En solicitudes posteriores, como cuando el usuario agrega productos al carrito de compras, el navegador enviará la cookie en el encabezado Cookie con cada solicitud. Al recibir esta cookie, el servidor podrá identificar al usuario mediante el sessionID y mantener su sesión activa, lo que permite recordar los productos en el carrito o mantener la autenticación sin necesidad de que el usuario vuelva a ingresar su información.

Las cookies pueden tener una serie de parámetros que determinan su comportamiento y cómo deben ser gestionadas. Los parámetros más comunes son:

- Dominio: El dominio de una cookie restringe su acceso y lectura únicamente a aquel que la haya establecido, lo que evita que sitios web ajenos puedan acceder a la información almacenada en el navegador del usuario.
- Path: Cada cookie está asociada a una ruta específica en el servidor. Esto significa que solo se enviará con las solicitudes que estén dentro de esa ruta definida, restringiendo su uso a áreas específicas de la página web.
- Tiempo de vida (Expires/Max-Age): Las cookies pueden tener un tiempo de vida limitado. Si se establece un valor Expires, la cookie se eliminará cuando esa fecha llegue. Si se utiliza Max-Age, la cookie se eliminará después de un periodo determinado en segundos desde que fue establecida. Dependiendo de este tiempo de vida, las cookies pueden clasificarse en dos tipos:
    - Cookies de sesión: Son temporales y se eliminan automáticamente cuando se cierra el navegador. Generalmente, se utilizan para mantener sesiones activas durante la navegación.
    - Cookies persistentes: Se almacenan en el dispositivo del usuario durante un periodo de tiempo más prolongado incluso después de cerrar el navegador. Estas cookies permiten que el sitio recuerde al usuario entre visitas.

Existen también restricciones de seguridad para proteger la privacidad y la integridad de las cookies. Algunas de las más importantes son:

- Secure: Una cookie marcada con el atributo Secure solo se enviará a través de conexiones HTTPS, lo que asegura que no se transmita por una conexión no segura, protegiendo su contenido.
- SameSite: Este atributo restringe el envío de cookies en solicitudes realizadas desde sitios externos. Existen tres opciones para este atributo:
    - Strict: La cookie solo se enviará en solicitudes realizadas desde el mismo sitio web que la ha generado.
    - Lax: Permite que la cookie se envíe en solicitudes de navegación de primer nivel, como cuando el usuario navega directamente a un sitio web.
    - None: La cookie se enviará en todas las solicitudes, independientemente del origen, aunque requiere que la cookie esté marcada como Secure (usada solo en HTTPS).

3. Cookies de terceros y su papel en la publicidad online

Las cookies de origen son creadas por la página web que estamos visitando y tienen como objetivo garantizar una experiencia de usuario fluida. Algunos ejemplos son la cookie de bienvenida, que nos permite iniciar sesión en un sitio web; la cookie del carrito de compra, que recuerda los productos que contiene; y la cookie de recomendación, que hace recomendaciones de productos basadas en nuestras preferencias.

Sin embargo, las cookies de terceros no son creadas por el sitio web que estamos visitando, sino por dominios que suelen estar asociados a servicios externos, como Google Ads. A diferencia de las cookies de origen, estas están destinadas principalmente a recordar preferencias y rastrear nuestro comportamiento en línea a través de diferentes sitios web. Hoy en día, su uso se ha convertido en una herramienta fundamental para la publicidad online, ya que permiten construir perfiles detallados de los usuarios, lo que facilita la segmentación y la personalización de los anuncios.

Un ejemplo de cómo funcionan las cookies de terceros en la publicidad es el caso de DoubleClick, una de las primeras plataformas publicitarias que aprovechó estas cookies para rastrear a los usuarios. DoubleClick utiliza imágenes de 1x1 pixel, también conocidas como “web beacons” o “balizas web”, que son pequeños archivos invisibles que se insertan en las páginas web o en los correos electrónicos. Además, se pueden incluso ocultar configurando la propiedad de visualización CSS en ninguno.

Estas imágenes tan pequeñas no interfieren con la experiencia del usuario, pero permiten rastrear su actividad. Cuando un usuario visita una página que contiene estas imágenes, su navegador envía información sobre la visita, incluyendo cookies de seguimiento. Este tipo de rastreo permite a los anunciantes saber qué sitios ha visitado el usuario, el tipo de navegador utilizado, además de la IP del dispositivo o la hora en la que se visitó la página, sin que este tenga conocimiento de ello.

Gracias a las cookies de terceros, la publicidad segmentada ha experimentado una auténtica revolución. Antes de su invención, los anuncios eran casi siempre generales y no tenían en cuenta las preferencias individuales de los usuarios. Sin embargo, con la capacidad de seguir la actividad de los usuarios en múltiples sitios web, las empresas comenzaron a ofrecer anuncios mucho más personalizados. Es por ello que cuando un usuario visita por ejemplo varias páginas sobre ropa, es probable que luego vea anuncios relacionados con ello en su próximo sitio web. Con la creación de estas cookies, los anunciantes pudieron seguir a los usuarios a través de diferentes sitios web, creando perfiles detallados sobre sus intereses, preferencias y hábitos de compra, pudiendo incluso monetizar su comportamiento.

Sin embargo, este cambio también trajo consigo una serie de problemas éticos y legales. El más polémico siendo la violación de la privacidad. Todo esto llevó a la creación de distintas leyes y reglamentos, para que las páginas web obtengan el consentimiento explícito antes de almacenar cookies en el dispositivo.

4. El debate ético y legal: Regulaciones y su impacto

Las cookies son objeto de debate ético y legal debido a su capacidad para rastrear y registrar las actividades en línea de los usuarios. En este contexto, han surgido normativas como el Reglamento General de Protección de Datos (GDPR) en Europa y la Ley de Privacidad del Consumidor de California (CCPA), cuyo principal objetivo es proteger la privacidad de los usuarios y proporcionar un mayor control sobre sus datos personales.

El GDPR entró en vigor en 2018 y exige que las empresas obtengan el consentimiento explícito de los usuarios antes de almacenar o procesar cookies que no sean estrictamente necesarias para el funcionamiento del sitio web. De manera similar, la CCPA exige que las empresas informen a los usuarios sobre los datos que recopilan y les proporcionen la oportunidad de optar por no vender su información personal. Estas políticas tienen como objetivo empoderar a los usuarios, pero su implementación ha tenido resultados mixtos.

La distribución de banners es uno de los resultados más visibles del cumplimiento normativo. Sin embargo, estos han sido criticados por sus diseños confusos y poco claros, que a menudo priorizan una aceptación rápida. Las investigaciones muestran que la mayoría de los usuarios simplemente los aceptan sin leerlos, lo que plantea dudas sobre la eficacia real del consentimiento informado. Este fenómeno, conocido como fatiga del clic, reduce significativamente el control del usuario.

Respecto al impacto en las empresas, cumplir con estas regulaciones es una tarea especialmente costosa para las pequeñas y medianas empresas, que a menudo carecen de los recursos técnicos y legales para implementar soluciones sólidas de gestión del consentimiento. Por su parte, las grandes corporaciones han desarrollado sofisticadas herramientas de gestión de la privacidad que les permiten seguir operando sin mayores inconvenientes, consolidando su dominio en el mercado digital.

Muchos expertos argumentan que el enfoque de la legislación es insuficiente, ya que el problema no radica en las cookies como herramienta, sino en cómo se utilizan y en la falta

de transparencia en torno a su propósito. Las regulaciones actuales, aunque bien intencionadas, se centran mayoritariamente en exigir el consentimiento de los usuarios, lo que genera una carga adicional para ellos sin abordar de manera efectiva el problema de fondo: la necesidad de un rediseño ético de los modelos de negocio que dependen del rastreo intensivo.

En resumen, aunque las regulaciones han mejorado la protección de la privacidad, no han eliminado por completo las prácticas de rastreo invasivo. Un enfoque más efectivo debería incluir educación para los usuarios, reglas técnicas claras y una supervisión activa para asegurar un uso ético de los datos.

5. Alternativas actuales a las cookies: Huellas digitales, Privacy Sandbox, …

Explicación de nuevas técnicas de rastreo

Actualmente se han desarrollado nuevas técnicas de rastreo en las webs como son las huellas digitales o los identificadores únicos basados en atributos del navegador.

En primer lugar, las huellas digitales identifican de manera única a un usuario mediante el análisis de características específicas del navegador y de su ordenador. Estas características incluyen datos como el idioma, resolución de pantalla, versión del navegador y sistema operativo, entre otros. Con esta información del usuario se crea un perfil único que permite a las webs seguir a los usuarios aunque éstos no utilicen cookies.

En segundo lugar, los identificadores únicos basados en atributos del navegador funcionan mediante la asignación de un código único generado a partir de datos del navegador o dispositivo, como por ejemplo el modelo, el software o las configuraciones específicas del usuario. A diferencia de las cookies, este mecanismo requiere que las webs mantengan las bases de datos centralizadas, para así poder asociar el identificador con la actividad del usuario, pero no dependen de ficheros almacenados localmente.

Introducción a la Privacy Sandbox de Google

Google ha desarrollado la Privacy Sandbox, la cual se basa en sustituir las cookies de terceros por otras tecnologías más centradas en la privacidad del usuario.

Por ello, se implementó FLoC (Federated Learning of Cohorts) con el fin de mejorar la privacidad a la hora de mostrar los anuncios al usuario. Consiste en un algoritmo de aprendizaje automático que agrupa conjuntos anónimos de usuarios por sus intereses y comportamiento. Más adelante este sistema de FLoC fue reemplazado por Topics API, este mecanismo se inspira en la publicidad basada en intereses, por lo que le muestra al usuario anuncios en relación a webs que visitó anteriormente.

Se pasó a utilizar Topics API ya que determina temas de interés sin tener que rastrear la

actividad de los usuarios, por lo que ofrece más transparencia y control sobre los datos ya que el usuario puede eliminar temas que no le gusten o inclusive desactivar Topics.
Además de Privacy Sandbox de Google, se implementaron otras herramientas como ITP (Intelligent Tracking Prevention) de Apple.

Comparativa con otras iniciativas (ITP - Apple)

Por su parte, Apple creó Intelligent Tracking Prevention (ITP) para combatir el seguimiento entre webs mediante el bloqueo de las cookies de terceros. De esta forma se elimina un gran número de las cookies de origen y se difuminan características específicas del usuario para dificultar su identificación.

En comparación, Google busca sustituir las cookies de terceros por Topics API para permitir compartir información general sin necesidad de un rastreo concreto, mientras que Apple mediante ITP limita las cookies de terceros y bloquea las huellas digitales, lo que dificulta el seguimiento del usuario entre webs.
Por ello, Privacy Sandbox se enfoca en un equilibrio entre privacidad y personalización, en cambio ITP prioriza la privacidad ya que evita cualquier intercambio de datos que no sea consentido por el usuario.

Críticas y limitaciones de estas nuevas tecnologías

En el caso de Privacy Sandbox de Google, ha recibido numerosas críticas por parte de los usuarios al no ser un avance en privacidad y ser únicamente un cambio de enfoque de lo que ya había. Esto se debe a que se continúa manteniendo un control bastante significativo sobre los datos de los usuarios ya que a pesar de la eliminación de las cookies de terceros, Topics API se basa en la personalización de los anuncios en relación a los intereses generales del usuario.

Por ello, estos mecanismos no son una solución definitiva de la privacidad de la web pero son un gran avance al poder limitar las cookies de terceros. Las críticas recibidas son en gran parte debido a que los cambios están más orientados a que las grandes compañías mantengan el control sobre la información de los usuarios en vez de proporcionar una verdadera privacidad.

6. Un futuro post-cookies: El camino hacia una navegación más privada

El mundo digital avanza hacia un modelo en el que las cookies, especialmente las de terceros, serán cosa del pasado. Este cambio está impulsado por un movimiento global que busca priorizar la privacidad de los usuarios y limitar el seguimiento invasivo. Empresas tecnológicas, gobiernos y usuarios están presionando por un ecosistema digital más ético y transparente, lo que plantea tanto oportunidades como desafíos para la industria. Google, por ejemplo, está eliminando progresivamente las cookies de terceros en Chrome, siguiendo el camino de Apple, cuyo navegador Safari ya bloquea estos rastreadores. Este cambio responde a la creciente demanda de privacidad, pero genera incertidumbre en la publicidad digital, que depende del rastreo para segmentar audiencias y medir resultados.

En respuesta, están surgiendo alternativas descentralizadas y herramientas de gestión de privacidad. Por ejemplo, el almacenamiento local de datos permite que los navegadores gestionen la información necesaria sin compartirla con servidores externos. Navegadores como Brave y Firefox ya ofrecen herramientas que bloquean rastreadores y permiten configuraciones avanzadas, poniendo al usuario en control de su privacidad.

Además, en el ámbito de la publicidad, se están desarrollando nuevas tecnologías para reemplazar las cookies de terceros mientras se mantiene la capacidad de segmentar audiencias. Entre estas propuestas destacan el Privacy Sandbox de Google, que incluye herramientas como Topics API para agrupar usuarios según intereses generales en lugar de rastrear su historial completo; los identificadores locales, gestionados directamente en los dispositivos de los usuarios; y la autenticación unificada, que asocia a los usuarios mediante un registro explícito en los servicios que utilizan, reduciendo la necesidad de cookies para identificarlos. Sin embargo, estas tecnologías enfrentan críticas por seguir favoreciendo a las grandes plataformas tecnológicas.

Este cambio tendrá un impacto significativo en el marketing digital, la experiencia de usuario y el diseño web. El marketing digital deberá apoyarse más en datos propios obtenidos directamente de los usuarios, priorizando su confianza y fidelidad, lo que complicaría la medición de campañas en múltiples plataformas. En cuanto al diseño web, la experiencia podría ser menos personalizada al principio, pero más limpia y libre de interrupciones intrusivas a largo plazo.

El camino hacia un futuro post-cookies no está exento de desafíos. Algunos escenarios posibles incluyen una mayor consolidación de poder en grandes plataformas como Google y Facebook, que podrían incrementar su dominio en el mercado publicitario al contar con vastos ecosistemas de datos propios. Además, las nuevas soluciones requieren una infraestructura avanzada, lo que podría ser costoso de implementar para pequeñas empresas. En conclusión, el futuro post-cookies promete una navegación más respetuosa con la privacidad, pero su éxito dependerá de la colaboración entre empresas, reguladores y usuarios para construir un ecosistema más ético y transparente.

Bibliografía

https://mailchi.mp/bonillaware/cookies

https://taggrs.io/es/first-party-vs-third-party-cookies/

https://www.aepd.es/guias/estudio-fingerprinting-huella-digital.pdf

https://hogantechs.com/es/baliza-web-interfaz-del-navegador-javascript-baliza-web/

https://blog.hubspot.es/marketing/que-es-floc#que

https://www.singular.net/blog/google-apple-privacy-marketing-attribution/