

Autor del Treball

Roger Rovira
rogerrov2006@gmail.com
@rovi_rover en X
Catalunya, Barcelona

Institució Acadèmica

INS. Lluís de Requesens
Tutor: Albert Flo
Curs: 2n de batxillerat
Any Acadèmic: 2023-24

Solitude Search

Preservant la Privacitat Digital.

Mitjançant Homomorphic Encryption i Private Information Retrieval, la plataforma dissenyada permet l'accés confidencial a una gran selecció d'articles.

Al GitHub

**Projecte**

Nom: Solitude
Autor: Roger Rora
Usuari: Gasofa06
Enllaç: github.com/Gasofa06/Solitude

Llenguatges: Rust, JavaScript, CSS, HTML i Python

Data de Creació: 18 d'agost de 2023

Etiquetes: homomorphic-encryption, private-information-retrieval, security, webapp

El codi de la plataforma està sotmès a la llicència MIT, la qual atorga el dret d'utilitzar, copiar, modificar, fusionar, publicar, distribuir, sublllicenciar i/o vendre còpies del programari, sempre que es compleixin les condicions específicades en: github.com/Gasofa06/Solitude/blame/main/LICENSE



**Solitude Search:
Preservant la Privacitat dels Usuaris**

Treball de Recerca del curs 2023-2024, subjecte a les condicions esmentades a baix.

Publicat el 14 de desembre de 2023.

**Disseny de Portada,
Redacció i Maquetació,**

Roger Rovira
(@rovi_roger en X)

**Termes de la Llicència
de Copyright**

Es permet compartir aquest document sempre que no es modifiqui el seu contingut en cap circumstància. Queda prohibit obtenir benefici de la venda de qualsevol material o element del document, així com de la publicació d'aquest, sense el permís previ per escrit de l'autor. En cas d'utilitzar el document o algun dels seus materials per a un propòsit sense ànim de lucre, s'haurà de citar adequadament el document.

Abstract

RESUM EN CATALÀ.

Paraules Claves:

Privacitat de les Dades

Centralització Seguretat

Dades

Homomorphic Encryption

Recuperació d'Informació Privada

En els darrers anys, les dades han adquirit un alt valor econòmic, impulsant a les empreses a maximitzar la recopilació d'informació fins al punt d'inferir dràsticament en la privacitat dels usuaris. Malgrat l'establiment de nous reglaments envers la protecció de dades i la implementació generalitzada de sistemes criptogràfics tradicionals, el dilema de la privacitat encara no ha cessat.

Tanmateix, com a tecnologia decisiva per a la solució d'aquest problema, l'*Homomorphic Encryption* s'ha convertit en una àrea d'investigació destacada en matèria de seguretat de la informació. Aquesta tecnologia permet efectuar operacions sobre missatges xifrats sense necessitat de desxifrar-los, essent l'únic sistema criptogràfic capaç d'assolir-ho.

El treball que prossegueix presenta els reptes relatius a la privadesa i la centralització de dades, a més d'explicar els fonaments de la plataforma dissenyada, anomenada Solitude Search. La qual esdevé com un motor de cerca de bases de dades dissenyat per garantir la privadesa dels usuaris. Per tal d'assolir la confidencialitat en les cerques, la plataforma estableix un sistema de Recuperació d'Informació Privada mitjançant l'ús d'*Homomorphic Encryption*, així que les sol·licituds enviades pels usuaris es xifrin de manera que ni el mateix servidor ni una empresa terciaria puguin arribar a conèixer el seu comportament de navegació.

Per implementar el sistema criptogràfic, s'ha utilitzat el codi obert de *Spiral Privacy Clients* proporcionat per l'empresa Blyss. Aquest consta de dos esquemes basats en el mètode de *learning with errors*: el sistema Regev i el Gentry-Sahai-Waters. Permetent que Solitude Search esdevingui capaç de protegir les dades dels usuaris en un panorama digital on la privacitat comença a ser una necessitat.

RESUMEN EN CASTELLANO.

Palabras Claves:

Privacidad de los Datos

Centralización Seguridad

Datos

Homomorphic Encryption

Recuperación de Información Privada

En los últimos años, los datos han adquirido un alto valor económico, impulsando a las empresas a maximizar la recopilación de información hasta el punto de afectar drásticamente en la privacidad de los usuarios. A pesar del establecimiento de nuevos reglamentos en cuanto a la protección de datos y la implementación generalizada de sistemas criptográficos tradicionales, el dilema de la privacidad aún no ha cesado.

Sin embargo, como tecnología crucial para la solución de este problema, la *Homomorphic Encryption* se ha convertido en un área destacada de investigación en seguridad de la información. Esta tecnología permite realizar operaciones sobre mensajes cifrados sin necesidad de descifrarlos, siendo el único sistema criptográfico capaz de lograrlo.

El trabajo que prosigue presenta los retos relacionados con la privacidad y la centralización de datos, además de explicar los fundamentos de la plataforma diseñada, llamada Solitude Search. La cual se deviene como un motor de búsqueda de bases de datos diseñado para garantizar la privacidad de los usuarios. Para lograr la confidencialidad en las búsquedas, la plataforma establece un sistema de Recuperación de Información Privada mediante el uso de *Homomorphic Encryption*, de modo que las solicitudes enviadas por los usuarios se cifren de manera que ni el propio servidor ni una empresa terciaria puedan llegar a conocer su comportamiento de navegación.

Para implementar el sistema criptográfico, se ha utilizado el código abierto de *Spiral Privacy Clients* proporcionado por la empresa Blyss. Este consta de dos esquemas basados en el método de *learning with errors*: el sistema Regev y el Gentry-Sahai-Waters. Permitiendo que Solitude Search sea capaz de proteger los datos de los usuarios en un panorama digital donde la privacidad comienza a ser una necesidad.

ABSTRACT IN ENGLISH.

Keywords:

Data Privacy

Centralization

Security

Data

Homomorphic Encryption

Private Information Retrieval

In recent years, data has acquired significant economic value, prompting companies to maximize the collection of information to the extent that it drastically impacts user privacy. Despite the establishment of new regulations regarding data protection and the widespread implementation of traditional cryptographic systems, the privacy dilemma has not ceased.

However, as a key technology for solving this problem, Homomorphic Encryption has become a crucial area of research in data security. This technology allows operations to be performed on encrypted messages without the need to decrypt them, making it the only cryptographic system capable of achieving this.

The work that follows presents challenges related to privacy and data centralization, in addition to explaining the foundations of the designed platform, named Solitude Search. It becomes a search engine for databases designed to ensure user privacy. To achieve confidentiality in searches, the platform establishes a Private Information Retrieval system using Homomorphic Encryption, so that requests sent by users are encrypted in a way that neither the server itself nor a third-party company can learn about their browsing behavior.

To implement the cryptographic system, the open-source code of *Spiral Privacy Clients* provided by the company Blyss has been used. This consists of two schemes based in the learning with errors method: the Regev system and the Gentry-Sahai-Waters system. Allowing Solitude Search to be able to protect user data in a digital landscape where privacy is becoming a necessity.

RÉSUMÉ EN FRANÇAIS.

Mots Clés:

Confidentialité des données

Centralisation

Sécurité

Données

Homomorphic Encryption

Récupération d'informations privées

Ces dernières années, les données ont acquis une valeur économique élevée, incitant les entreprises à maximiser la collecte d'informations au point d'en déduire drastiquement la vie privée des utilisateurs. Malgré l'établissement de nouvelles réglementations en matière de protection des données et la mise en œuvre généralisée des systèmes cryptographiques traditionnels, le dilemme en matière de confidentialité n'a pas encore disparu.

Cependant, en tant que technologie décisive pour résoudre ce problème, le *Homomorphic Encryption* est devenu un domaine de recherche important en matière de sécurité de l'information. Cette technologie permet d'effectuer des opérations sur des messages cryptés sans avoir besoin de les déchiffrer, étant le seul système cryptographique capable d'y parvenir.

Le travail qui suit présente les défis liés à la confidentialité et à la centralisation des données, ainsi qu'expliquant les fondamentaux de la plateforme conçue, appelée Solitude Search. Ce qui devient comme un moteur de recherche de base de données conçu pour garantir la confidentialité des utilisateurs. Afin d'assurer la confidentialité des recherches, la plateforme établit un système de Récupération d'Informations Privées grâce à l'utilisation du *Homomorphic Encryption*, de sorte que les demandes envoyées par les utilisateurs soient cryptées afin que ni le serveur lui-même ni une société tierce ne puissent connaître votre comportement de navigation.

Pour mettre en œuvre le système cryptographique, le code ouvert de *Spiral Privacy Clients* fourni par la société Blyss a été utilisé. Il s'agit de deux schémas basés sur la méthode du *learning with errors*: le système Regev et le Gentry-Sahai-Waters. Permettre à Solitude Search de devenir capable de protéger les données des utilisateurs dans un paysage numérique où la confidentialité devient une nécessité.

AGRAÏMENTS.

El treball que es presenta a continuació pretén ser un homenatge a totes aquelles persones que en un moment donat m'he anat trobant pel camí de l'enginyeria i la informàtica. Encara que potser no hem arribat a parlar sobre temes específics com l'*Homomorphic Encryption* i la recopilació de dades, les seves contribucions han estat indispensables en l'elaboració d'aquest projecte. Cada persona, des de la seva pròpia experiència i coneixement, ha enriquit el meu bagatge i ha afegit les dosis necessàries d'entusiasme per ajudar-me a progressar al llarg d'aquest trajecte.

Així doncs, m'agradaria començar expressant la meva gratitud al *Barcelona Supercomputing Center* (BSC) per oferir-me la possibilitat de participar en el programa de *Bojos per la Supercomputació*. Aquesta experiència no només m'ha proporcionat coneixements sinó que també ha despertat en mi un fort interès per continuar estudiant.

També vull agrair a la Sara Ibáñez, coordinadora del programa de *Bojos per la Supercomputació*, per fer-me d'enllaç i donar-me suport, així com als *Bojos* que han cursat aquest programa amb mi.

Agraeixo també l'assessorament d'en Dr. Leonardo Bautista, en Tarun Mohandas i en Mikel Cortes, tots tres investigadors del BSC en la recerca de *Blockchain Scalability and Sustainability*, per donar-me a conèixer la tecnologia en què es fonamenta aquest treball: l'*Homomorphic Encryption*.

A més a més, vull agrair a tots aquells companys i professorat que, d'una manera o una altra, m'han ajudat i motivat a continuar treballant.

I ja per acabar, m'agradaria expressar el meu agraïment a la meva família, per la seva infinita paciència i dedicació; i especialment, als meus pares, per suportar i recolzar totes les idees que, en el meu dia a dia, sorgeixen com del no-res.

Moltes gràcies a tots, aquest treball no hauria estat possible sense vosaltres.

Figura X.1: Imatge del Barcelona Supercomputing Center, que actualment compta amb el *MareNostrum 5 ACC*, el vuitè superordinador més potent del món i el tercer d'Europa.

Font: La imatge que es mostra ha estat extreta de la [pàgina web oficial del BSC](#).



“ “

[La implementació d’]un sistema que no posa en control als usuaris serà, immediatament o amb el temps, rebutjat per tants d’ells que no pugui esdevenir i continuar essent una tecnologia unificadora.

~ Kim Cameron
The Laws of Identity (2005, maig 11)

Traducció pròpia de l’anglès al català. Cita original: “A system that does not put users in control will, immediately or over time, be rejected by enough of them that it cannot become and remain a unifying technology”.

‘ ’

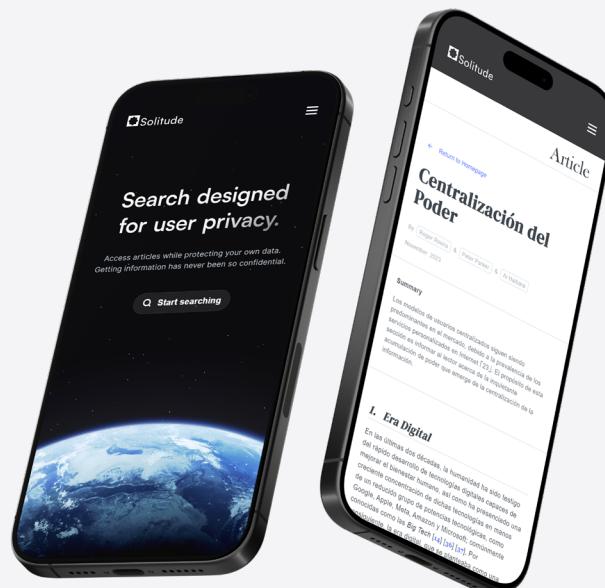
CONTINGUTS

| | | |
|---|--|--|
| Titul de la secció | Subapartats | |
| Presentació | Introducció 10 Objectius del Treball de Recerca 11 Metodologia 12 Motivació 13 Dificultats Trobades 13 | |
| Titul de la secció | Subapartats | |
| Plataforma Creada: Solitude Search | Breu Explicació de la Plataforma 37 Eines Emprades 38 Programació 40 Disseny de la Pàgina Inicial 44 Disseny del Cercador 46 | |
| Titul de la secció | Subapartats | |
| Centralització de Dades | Data Monopolies 15 Dades Personals 16 El Valor Ocult de les Dades 17 Increment d'Informació 19 Facebook i Cambridge Analytica 20 Data Breaches 21 Google, Dont Be Evil 22 Algunes de les Aplicacions de Google 24 Ingenuïtat Humana 25 | |
| Titul de la secció | Subapartats | |
| Conclusions Finals del Projecte | Conclusió 48 | |
| Titul de la secció | Subapartats | |
| Solucions per a la Privacitat | Violacions de Privacitat 27 Solucions Jurídiques 28 Confiança en Abundància 29 Declaracions de Privacitat 30 Solucions Tecnològiques 31 Problemes amb la Criptografia 32 El Sant Grial de la Criptografia 34 | |
| Titul de la secció | Subapartats | |
| Referències | Articles 50 Documents Legals 51 Pàgines Web 51 Articles Periodístics en Línia 52 Material Audiovisual 52 Comunicats 52 Llibres 52 | |

Presentació del Projecte

Presentació del Projecte

Aquest treball de recerca pretén demostrar la viabilitat d'una solució a la manca de privacitat a Internet, mitjançant la creació d'una plataforma completament privada. La secció de continuació serveix com a introducció a la plataforma i al treball en general.



INTRODUCCIÓ.

És difícil mantenir secrets en l'era digital. Malgrat la criptografia i les diverses legislacions implementades, el problema de la privacitat continua assetjant als usuaris en Internet.

Amb cada missatge, cada cerca i cada interacció; les empreses aprenen quelcom més d'informació sobre nosaltres. A mesura que aquestes empreses obtenen més informació, més poderoses esdevenen i més vulnerables ens tornem nosaltres.

Avui en dia, més de 5000 milions de persones arreu del món estan interconnectades a través d'Internet per tal d'aconseguir o enviar informació. No obstant això, a mesura que interactuem amb les diverses plataformes que usem, aquestes recopilen dades sobre nosaltres. I, cada vegada més, el preu que hem de pagar per estar connectats és la nostra privacitat [Kov12]. Encara que a molts ens agradaria creure que plataformes com Google Search, Instagram i TikTok són d'accés gratuït, en la pràctica, és ben sabut que no ho són; vist que utilitzen les nostres dades com a forma de compensació.

Si bé s'espera que siguem els propietaris de la nostra informació, per poder participar mínimament en la vida moderna, ens veiem obligats a renunciar a aquest dret fonamental. Degut, en part, a la proliferació de diverses empreses tecnològiques que assumeixen el control de les nostres dades, amb l'únic objectiu d'obtenir-ne benefici.

Tanmateix, la implementació d'un sistema que no posa en control als usuaris serà, immediatament o amb el temps, rebutjat per tants d'ells que no pugui esdevenir i continuar essent una tecnologia unificadora [Cam05]. Així doncs, per tal de preservar la sostenibilitat en Internet, és imperatiu que el poder retrocedeixi de l'empresa a l'individu i que aquest últim assoleixin la privacitat que li pertoca.

OBJECTIUS DEL TREBALL DE RECERCA.

Privacitat (objectiu general)

L'objectiu general del projecte és analitzar els problemes de la privacitat en l'era digital i comprendre per què sorgeixen, amb la finalitat de dissenyar una plataforma que garanteixi una confidencialitat màxima, mitjançant la tecnologia.



Centralització

Durant l'escrit, es pretén examinar l'estructura de l'era digital i identificar les conseqüències de la centralització de les dades, així com il·lustrar-ho amb casos reals de filtracions d'informació i situacions de monopolis comercials basats en les dades.

→ Llegeix més en la [Secció 1](#).

HE/PIR

La part teòrica té, entre altres, el propòsit d'introduir el concepte d'*Homomorphic Encryption*, una tècnica criptogràfica que assegura una confidencialitat total, i aprofundir en les seves aplicacions, posant èmfasi en els usos de Recuperació d'Informació Privada (*Private Information Retrieval*, PIR).



Branding

Al llarg del desenvolupament del projecte, s'ha buscat establir una identitat de marca mitjançant la creació del logotip i la definició de la paleta de colors, entre altres elements.

Problemes

La part escrita també té l'objectiu d'examinar la problemàtica de la privacitat en l'era de la informació, aprofundir en les solucions plantejades des d'una perspectiva legal i tecnològica, i il·lustrar les restriccions d'aquestes solucions per abordar de manera integral la manca de privacitat dels usuaris.



Programació

Pel que fa a la part pràctica, el projecte pretén desenvolupar un dels primers navegadors de bases de dades que preservi de manera completa la privacitat dels usuaris, mitjançant algoritmes d'*Homomorphic Encryption*.

Això implica també dissenyar la interfície, programar l'estructura interna del servidor, implementar un autocompletat eficient en el costat del client, minimitzar la transferència de dades i integrar una *reverse proxy*, així com altres característiques que es detallaran més endavant.

→ Llegeix més en la [Secció 3](#).

UX/UI

En el desenvolupament de la plataforma, Solitude Search, s'ha procurat crear una interfície d'usuari amb un disseny atractiu i intuitiu, assegurant-se al mateix temps que mantingui un bon rendiment i una bona accessibilitat.

Disseny de l'article

En la maquetació de l'article, s'ha procurat que la seva presentació sigui elegant i atractiva. Amb aquesta finalitat, s'han creat gràfics, il·lustracions i disposicions utilitzant diverses aplicacions d'Adobe.

METODOLOGIA.

El treball que segueix està estructurat en tres seccions. D'entrada, s'analitzen els diferents problemes que deriven de la centralització, oferint exemples de casos de monopolis basats en les dades i explicitant com aquests impacten en la nostra privacitat i el nostre lliure albir. Aquesta informació s'extreu d'articles científics i documents financers, com els informes 10-K.

La segona secció presenta les solucions que s'han proposat per abordar la manca de privadesa en l'era digital. S'exploren els mètodes legals i tecnològics, es detallen les seves limitacions, i s'introduceix un concepte criptogràfic que pot canviar les regles del joc. Com en el cas anterior, aquesta secció s'ha basat en diversos articles científics i documents legals.

En darrer lloc, la tercera secció tracta la dimensió pràctica del treball, que consisteix a desenvolupar un cercador que garanteixi la confidencialitat de l'usuari mitjançant proves criptogràfiques, per tal d'ofrir una solució específica a la manca de privacitat en Internet. En aquesta secció, es detallaran els fonaments que regeixen a Solitude Search, juntament amb els seus avantatges, limitacions i singularitats. La plataforma s'ha dissenyat utilitzant diversos llenguatges de programació com Rust, JavaScript, HTML, CSS i Python, junt amb projectes de codi obert com Spiral Privacy Clients, Actix Web i NGINX. També s'han emprat tecnologies com WebAssembly i Homomorphic Encryption, així com diverses estructures de dades, incloent-hi un Nested Hierarchical Tree i un Trie per a la cerca dels títols dels articles des del dispositiu del client.

Al llarg del treball, se segueix una rigorosa metodologia científica, citant les referències consultades en tot moment i incorporant elements gràfics creats amb aplicacions d'Adobe, com Photoshop i Adobe Illustrator.

A més a més, l'editor de text utilitzat per a la creació de l'article ha estat Adobe InDesign, el qual proporciona una gran flexibilitat a l'hora de dissenyar l'aspecte de l'article, tot i que també n'afegeix dificultats.

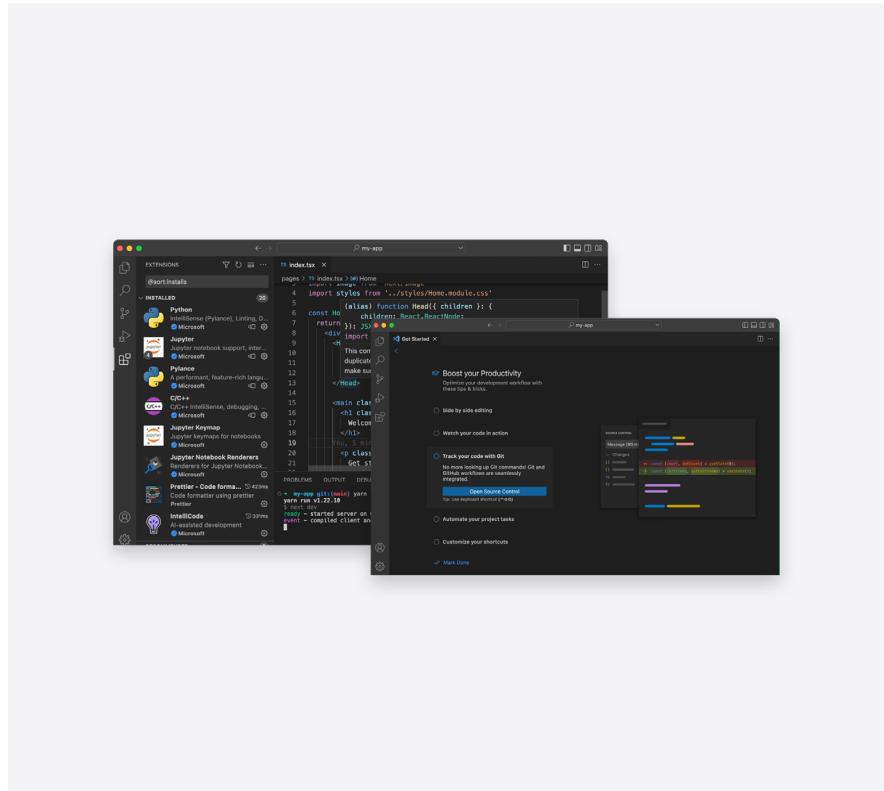


Figura 0.1: L'entorn integrat de desenvolupament utilitzat (IDE, per les seves sigles en anglès) ha estat el Visual Studio Code. El porto fent servir des de fa anys i considero que és el millor, amb un gran abast d'extensions que permeten adaptar l'entorn al projecte i facilitar-ne el desenvolupament.

Font: La imatge s'ha extret de la [pàgina web espanyola de Visual Studio](#).

MOTIVACIÓ.

El món ha experimentat canvis constants al llarg de la història, molt abans que jo nasqués i abans que tots els qui ara habiten la Terra ho fessin. No obstant això, m'aventuro a afirmar que mai havia ocorregut una acceleració tecnològica d'aquesta magnitud en la història. En temps de canvis com aquest és quan hem de garantir que s'estableixin els equilibris adequats. Per això, he decidit centrar aquest treball en la cerca d'una solució al desafiament més urgent i important que amenaça l'era digital: el problema de la privadesa.

Crec fermament en la importància d'equilibrar l'avanç tecnològic amb la protecció dels nostres drets i llibertats. El treball realitzat reflecteix aquesta convicció meva, en la qual hauríem de preservar un valor fonamental que ens permeti recuperar el control de la nostra informació personal en la societat que està per venir.

LA PRIVACITAT NO ÉS UNA OPCIÓ.

La privacitat és la capacitat de poder decidir a qui diem què, de limitar qui té accés a les nostres comunicacions i la nostra informació. La privacitat és respecte; si algú té un secret que vol mantenir en privat, és irrespectuós desobeyir el seu desig. La privacitat és un dret i res més.



Figura 0.2: La privacitat no és una opció i no hauria de ser el preu que hem de pagar per només estar connectats a Internet.

Font: S'ha dissenyat amb una imatge extreta de la [pàgina web de El Corte Inglés](#) i prenen inspiració d'un dels anuncis d'Apple sobre la privacitat.

DIFÍCULTATS TROBADES.

Durant gairebé una desena de mesos en què s'ha anat elaborant el projecte, hi ha hagut diverses dificultats que, en aquell moment, semblaven insuperables.

En primer lloc, una de les principals complicacions del projecte ha estat la implementació de l'*Homomorphic Encryption*, ja que el nombre d'aplicacions i documentació disponibles són certament limitats, a més de la seva inherent complexitat.

En segon lloc, la configuració del servidor Actix i la *reverse proxy* també han generat alguns maldecaps, car fins ara només havia programat servidors que implicaven pocs processos, a diferència d'aquest projecte, que requereix una infraestructura més complexa per permetre l'accés a informació privada.

La tercera gran complicació, si no és que n'oblido alguna, ha estat el disseny de l'article. Mai havia utilitzat InDesign, i va ser tot un repte aprendre a utilitzar-ho; fent i desfent canvis fins a arribar a un disseny agradable.

Si bé aquests són les tres complicacions que se'n destaquen, la part pràctica ha comportat mil i una dificultats més, atès que la programació implica resoldre problema rere problema fins a arribar a l'objectiu final.

Centralització de Dades

Centralització de Dades



Els models d'usuaris centralitzats continuen sent predominants en el mercat, a causa de la prevalença dels serveis personalitzats en Internet. El propòsit d'aquesta secció és informar al lector sobre la inquietant acumulació de poder que emergeix de la centralització de la informació.

DATA MONOPOLIES.

En les últimes dues dècades, la humanitat ha estat testimoni del ràpid desenvolupament de tecnologies digitals capaces de millorar el benestar humà. Així com ha presenciat una creixent concentració d'aquestes tecnologies en mans d'un reduït grup de potències tecnològiques, com Google, Apple, Meta, Amazon i Microsoft; sovint conegudes com les *Big Tech* [McI19] [Stu18] [BC21]. Per consegüent, l'era digital, que es plantejava com una excel·lent oportunitat per democratitzar les institucions i millorar l'accés a la informació i al coneixement [MTM07] [McI19], ha resultat ser una època contradictòria al que s'esperava. En lloc d'alliberar a la humanitat i fomentar la seva independència, l'era digital ha sotmès a les persones a la simple mercè de les grans companyies tecnològiques.

Fa vint-i-cinc anys, Meta no existia, Google era tan sols un projecte universitari i Amazon només s'enfocava en la venda de llibres [McI19]. Pro, tot i això, són actualment considerades com algunes de les empreses més valuoses i influents del món. Donada la seva immensa magnitud i abast, aquestes companyies tenen la capacitat d'influir en el comportament dels usuaris i l'economia internacional [BC21]; atorgant-los un poder sense precedents que es manifesta en diversos aspectes de la nostra vida quotidiana.

Un exemple rellevant és el cas de Cambridge Analytica, en el qual es va posar de manifest la funció que exerciren les dades acumulades de Facebook en les cincantenes octaves eleccions presidencials dels Estats Units. Demostrant la gran repercussió que les dades tenen sobre l'opinió pública i els processos democràtics. De la mateixa manera que Meta s'ha apoderat de la informació dels seus usuaris, tant Google com altres companyies han estat capaces de monopolitzar la informació de diverses maneres [McI19], col·locant-se en una posició privilegiada per capitalitzar la innovació futura i oferir serveis personalitzats.

DADES PERSONALS.

Les *Big Tech* han aconseguit establir-se entre les companyies més rellevants i poderoses del món gràcies a la seva innata habilitat per recopilar i utilitzar grans quantitats de dades en benefici del seu propi creixement. Google, per exemple, compta amb una àmplia base de dades que li permet identificar patrons i tendències per millorar els resultats de cerca que ofereix als seus usuaris [FM12] [McI19] [HH13]; i, de manera similar, Meta personalitza el contingut que mostra als seus consumidors basant-se en els interessos i preferències individuals [BC21].

Aquests escenaris comporten que els consumidors se sentin limitats a fer servir un servei específic en comptes de la seva opció preferida, a causa de la comoditat i adaptabilitat de la plataforma en qüestió. Optant d'aquesta manera per no emprar un servei que els podria proporcionar una major privadesa [McI19] [Stu18].

Nota 1: DuckDuckGo és una empresa independent que s'esforça a proporcionar privacitat a tothom a Internet. És coneguda principalment pel seu motor de cerca en línia, que es diferencia dels competidors per la seva política de recopilació de dades i les seves eines contra els rastrejadors.

Per exemple, els usuaris poden preferir les polítiques de privacitat de DuckDuckGo¹, però romandre amb el motor de cerca dominant (Google), que, tot beneficiant-se de la seva àmplia xarxa d'usuaris i de la recopilació de dades que efectua, ofereix millors resultats de cerca [HH13] [Stu18] [SG17]. De manera semblant, els conductors poden preferir un servei de navegació que prioritzi la confidencialitat, però quedar-se amb l'aplicació dominant de Google, coneguda com a Google Maps [Stu18]; quina finalitat no és només presentar la millor informació de trànsit, sinó que també busca recopilar la quantitat més gran de dades possible.

Si ens enfoquem en el sector dels motors de cerca, és evident que Google assoleix un rol predominant en la indústria [HH13] [KSh20]. Aquesta posició hegemònica es deu a la capacitat de Google d'analitzar les diverses dades que recopila amb els seus productes. Ja que així pot millorar els perfils d'usuari amb informació exclusiva que cap altre competidor posseeix. Aquest avantatge captiva els usuaris i els incentiva a utilitzar els serveis de Google, atès que no hi ha cap altra plataforma que ofereixi una personalització comparable o uns resultats de cerca tan precisos.

No obstant això, és important considerar si aquesta raó justifica plenament l'extensió en què Google recopila informació. Hem de comprometre les nostres dades personals per tal d'obtenir resultats més precisos? O és possible trobar un equilibri entre la comoditat i la protecció de dades?

| | Google | Bing | Yahoo! | DuckDuckGo |
|------|--------|------|--------|------------|
| 2010 | 90.91 | 3.46 | 3.93 | |
| 2014 | 89.81 | 3.63 | 3.57 | 0.04 |
| 2018 | 91.4 | 2.82 | 2.15 | 0.29 |
| 2020 | 92.07 | 3.19 | 1.36 | 0.69 |

Figura 1.1: Taula que conté la participació percentual del mercat dels motors de cerca a escala mundial.

Font: Elaboració pròpia amb informació extreta tant de *StatCounter* (Search Engine Market Share WorldWide) com d'una *publicació en Photutorial* de Matric Broz sobre les estadístiques d'usuaris de DuckDuckGo.

EL VALOR OCULT DE LES DADES.

Partint del poder que les empreses assoleixen amb les dades, s'infereix que la qualitat del producte està estretament relacionada amb la quantitat d'informació recopilada. Vist que aquesta permet a les companyies oferir contingut més precís.

Tanmateix, a causa de les dades que ens extreuen, les empreses són capaces d'identificar patrons, tendències i oportunitats de negoci [FM12] [McI19]; fet que els permet generar ingressos mitjançant models publicitaris o serveis addicionals. A més, la recopilació massiva d'informació atorga a les companyies la possibilitat de desendanar un efecte de xarxa basat en les dades personals, conegut en anglès com a *data network effect* [McI19] [HH13] [SG17].

Els *network effects* són tals que com més usuaris hi ha en una plataforma, més valiosa esdevé aquesta per a cada usuari [FM12] [Yoo12] [Stu18]. Els telèfons constitueixen un exemple clàssic. A mesura que més persones adquireixen dispositius telefònics, més persones estarán disponibles per realitzar trucades. Així mateix, tenir un major nombre de persones a les quals trucar augmenta el valor inherent de posseir un telèfon [Stu18].



Figura 1.2: El Nokia 6110, telèfon mòbil compatible amb la tecnologia GSM (*Global System for Mobile Communications*), va ser anunciat el 18 de desembre de 1997 i introduït al mercat durant l'any 1998.

Font: La imatge ha estat extreta de la [web](#) [It's Nice That](#).

En el context de les dades, en canvi, a mesura que la plataforma atreu més usuaris i acumula més informació, la seva habilitat per personalitzar el contingut s'intensifica i la comoditat de l'usuari augmenta. Aquest aspecte fomenta el retorn de l'usuari i estimula la voluntat de compartir informació personal. A més, com que ningú pot oferir serveis tan individualitzats, el negoci de l'empresa es consolida en el mercat.

En aquest sentit, Facebook recull informació dels seus usuaris per generar béns i serveis personalitzats [Stu18], amb l'objectiu últim d'augmentar la quantitat d'usuaris en la xarxa i la seva dependència d'ella, de manera que s'aconsegueixi bloquejar el nombre més gros d'usuaris en el sistema [FM12]. En efecte, a mesura que augmenta la quantitat d'usuaris en la xarxa i la plataforma obté més dades sobre nosaltres, es genera un major nivell de comoditat i una major dependència respecte al servei; la qual cosa contribueix al creixement i la consolidació de Facebook en el mercat de les xarxes socials.

Nota 2: Els ingressos nets anuals d'Alphabet (la companyia matriu de Google) per a l'any 2021 van ser de 76.033 mil millions de dòlars [Alp22]. Això implica que Google va pagar aproximadament a Apple una suma equivalent al 19,74% dels ingressos nets de la seva empresa matriu. A més, és probable que Google faci pagaments encara més grans a Apple per assegurar-se que Microsoft no els superi [Mor21].

Aquest mateix fenomen també es pot aplicar als motors de cerca [Stu18], així com a la computació en núvol i altres sectors [FM12]. Així és que diverses empreses tecnològiques ofereixen plataformes de comerç, xarxes socials o navegadors de manera *gratuïta*, pues utilitzen les nostres dades com a forma de pagament [Stu18].

Si les dades no tinguessin un valor elevat o no disposessin d'una importància transcendental, resultaria improbable que Google hagués pagat 82 milions de dòlars el 2009 a Apple, per assegurar la seva posició com a motor de cerca predeterminat en Safari [SG17]. De manera similar, els pagaments de 1000 milions de dòlars que Google va realitzar el 2013 i el 2014 pel mateix propòsit [SG17] no trobarien cap justificació; i resultaria encara més difícil comprendre el desemborsament de 15000 milions de dòlars que Google va realitzar el 2021 amb el mateix objectiu² [Mor21] [KSh20].

Així mateix, en cas que les dades no posseïssin un alt incentiu econòmic, resultaria difícil explicar per què Meta va pagar 16000 milions de dòlars en l'adquisició de WhatsApp, una empresa de només 60 empleats i sense actius tangibles [McI19]. Encara més, hi hauria poca o cap justificació amb la qual explicar l'eliminació de la petita tarifa de WhatsApp (0,99 dòlars) que hi havia imposta en alguns països [Stu18].

Malgrat que els usuaris (p. ex. de WhatsApp o Google) reben el benefici immediat d'utilitzar un servei gratuït, és important tenir en compte que el cost a llarg o curt termini de divulgar informació personal pot resultar fins i tot més gran que el de pagar una modesta tarifa anual. A causa de la manca de coneixement, per part dels usuaris, sobre com empraran les seves dades en el futur i per quines entitats seran fetes servir [Stu18].

SI NO HO ESTÀS PAGANT, NO ETS EL CLIENT; ETS EL PRODUCTE QUE ES VEN.

~ Andrew Lewis
MetaFilter: User-driven discontent (2010)



Figura 1.3: Cita d'Andrew Lewis, traducció pròpia del angle al català. Cita original: *If you are not paying for it, you're not the customer; you're the product being sold.* La cita s'ha extret de [MetaFilter](#).

Font: La imatge ha estat presa de la il·lustració fotogràfica de Natalie Matthews-Ramo, publicada a [Salte](#).

INCREMENT D'INFORMACIÓ.

Segons un informe publicat per IBM el 2013, el 90% de les dades existents fins a aquell moment havien estat generades en els dos anys anteriors [ZTL+18]; i, en una publicació més recent [IBM20], realitzada el 2020, IBM ha assenyalat que la quantitat de dades generades cada dia se situa al voltant dels 2.5 exabytes. Aquesta gran quantitat de dades ja no es limita únicament a informació bàsica (p. ex., nom, gènere, edat i correu electrònic), sinó que ara també recull una àmplia varietat de dades no estructurades. Entre les quals s'inclouen registres de navegació, dades de transaccions, arxius de correu electrònic, missatges de text, informació geoespatial, imatges, contactes, relacions i fins i tot preferències i opinions personals [McI19].

A partir d'aquest augment en la quantitat d'informació generada, va sorgir el concepte de *Big Data*. El qual es defineix com el conjunt de dades, tant estructurades com no estructurades, que es distingeixen per quatre característiques fonamentals conegeudes com les quatre Vs: volum, velocitat, varietat i valor.

En el context del *Big Data*, l'èmfasi no recau en la qualitat de les dades, sinó en la quantitat [McI19]. De manera que totes les companyies que tinguin accés al *Big Data*, així com la destresa per analitzar-ho, posseiran l'habilitat d'indagar tant en els interessos i inclinacions individuals (de cada persona) com en les tendències i preferències col·lectives (p. ex. d'un país, una ciutat o una comunitat); mitjançant el reconeixement de patrons i formes de comportament.

El fet de conèixer les preferències de cada individu o col·lectiu permet a les companyies influir en les accions i la presa de decisions dels seus usuaris; la finalitat d'influir en el nostre comportament pot ser merament econòmica o senzillament democràtica. Conseqüentment, aquesta capacitat comporta una major concentració del poder i una disminució en la diversitat del mercat; fet que resulta en una progressiva limitació de la privacitat i llibertat individual.

Tota la informació que les empreses obtenen sobre nosaltres es troba completament centralitzada i resguardada als seus propis servidors.

No obstant això, l'emmagatzematge centralitzat d'una quantitat tan gran de dades implica la irrupció de problemes derivats de la centralització, com ara l'eliminació intencionada de dades sensibles [Shr20], l'ús indegit d'informació confidencial, l'accés no autoritzat a dades personals o la manca de control dels usuaris sobre aquestes.

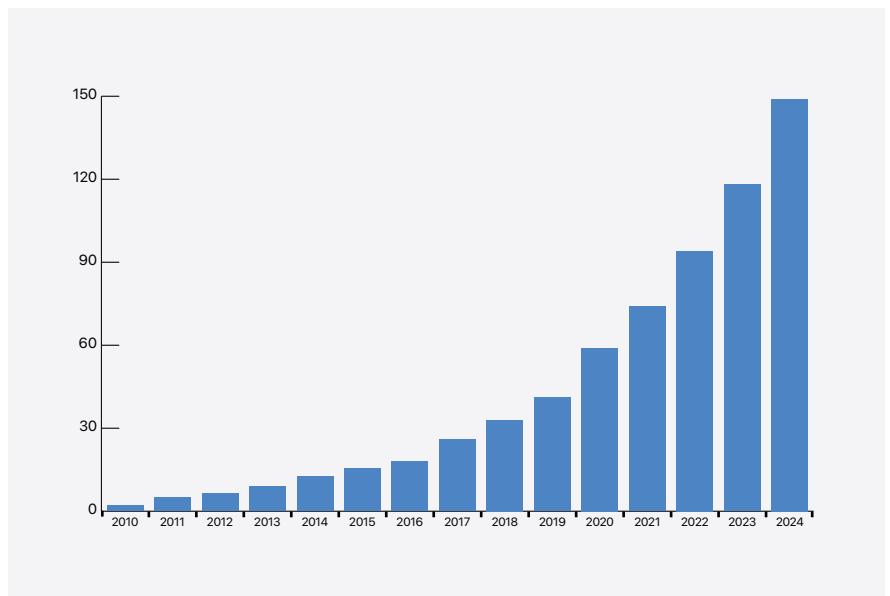


Figura 1.4: Volum de dades generades a escala mundial amb projeccions per al període de 2021-2024. Les dades estan expressades en zettabytes.

Font: Elaboració pròpria amb dades de *Exploding Topics*.

FACEBOOK I CAMBRIDGE ANALYTICA.

Escàndols, com el de Cambridge Analytica, han demostrat que hi ha deficiències significatives en els mecanismes de gestió i administració de dades personals [Gia20]. No més lluny, però, aquests esdeveniments també han destacat la importància de revisar i reforçar la protecció de privacitat dels usuaris en l'àmbit digital.

D'acord amb la informació proporcionada per Meta Investor Relations [Met23], Facebook actualment consta amb més de 2000 milions d'usuaris actius diàriament, el que equival aproximadament a un quart de la població mundial. Quan es treballa amb xifres d'aquesta magnitud, inclús la modificació més petita o la menor transgressió dels drets del consumidor poden ocasionar conseqüències catastròfiques. La rellevància d'aquestes conseqüències augmenta quan es tracta d'informació personal amb caràcter polític, com succeeix amb les dades recopilades per la plataforma Facebook.

CEO DE META PLATFORMS.

Mark Zuckerberg és un empresari i cofundador de Meta Platforms, abans coneguda com a Facebook, Inc. Ha estat el president executiu de l'empresa des de la seva creació. Malgrat la seva influència en la indústria tecnològica, Meta ha rebut crítiques significatives per diversos problemes de privacitat, generant preocupacions i debats sobre la seva gestió de dades.



Figura 1.5: Mark Zuckerberg, *chief executive officer* (CEO) de Meta Platforms. És la cara de l'empresa i el seu màxim responsable.

Font: La imatge s'ha extret d'[ONGPNG](#).

El 2018, es va formular una acusació contra Cambridge Analytica, una companyia privada especialitzada en ànalisi de dades, per haver utilitzat informació confidencial dels usuaris de Facebook per generar publicitat dirigida en suport a la campanya electoral de Donald Trump [RCC18]. Segons l'informe presentat per Mike Schroepfer [Sch18], que en aquell moment ocupava el càrrec de Chief Technology Officer (CTO) de Meta Platforms, Inc., es va revelar que Cambridge Analytica havia obtingut dades potencials de més de 87 milions d'usuaris de Facebook. L'obtenció d'aquestes dades es va realitzar mitjançant l'explotació d'una vulnerabilitat en les interfícies de programació d'aplicacions (APIs) de la plataforma. A més, cal destacar que l'apropiació de les dades en qüestió es va exercir sense el degut consentiment dels usuaris afectats.

Posteriorment, les dades van ser analitzades amb la finalitat de crear perfils psicològics i mostrar anuncis específics per atreure votants. En termes generals, la campanya va llançar 4000 anuncis publicitaris que, en conjunt, van aconseguir més de 1400 milions d'impressions [McI19]. S'argumenta que els anuncis dirigits elaborats per Cambridge Analytica van ser fonamentals per l'èxit de la campanya electoral de Trump en el 2016, ja que va permetre dirigir-se als votants amb missatges específics que ressonaven amb les seves necessitats més intimes i els seus desitjos més preuats.

El cas exposat demostra que, tot i assumir que Facebook no té gens d'interès polític, la seva funció com a entitat centralitzada i la seva habilitat per recopilar i emmagatzemar grans quantitats de dades, va originar la conseqüència no desitjada de permetre que tercers utilitzessin aquesta informació amb fins maliciósos.

DATA BREACHES.

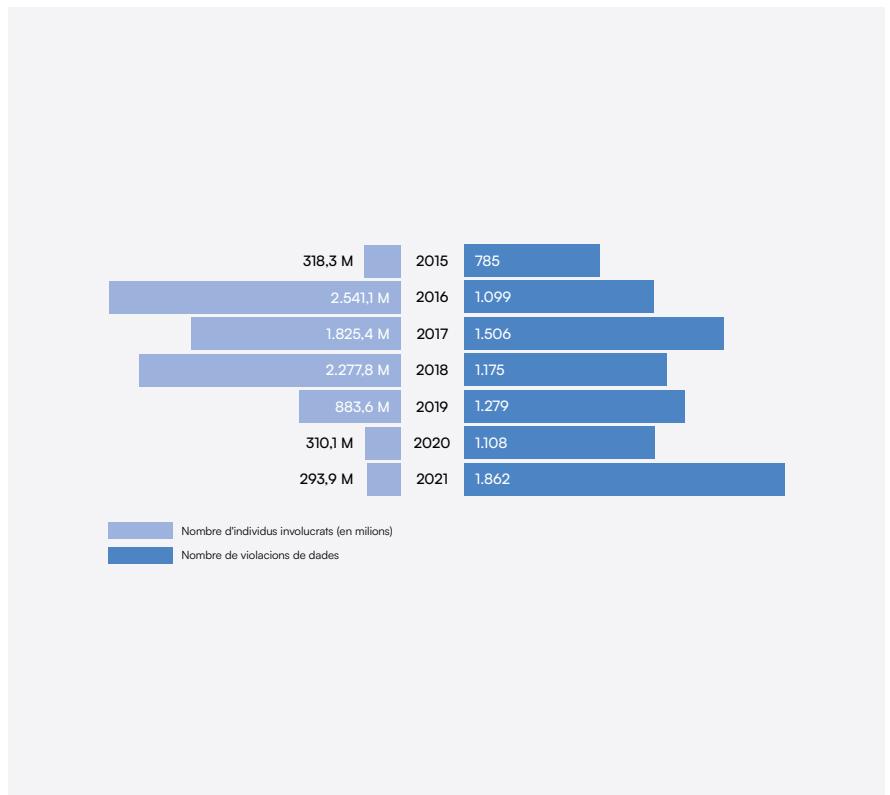
Les males gestions de Facebook i altres companyies tecnològiques que condueixen a grans filtracions d'informació confidencial es coneixen com a data breaches. Aquestes pràctiques, tot i que poden semblar esporàdiques, són força freqüents. Només als Estats Units, durant l'any 2022, es van registrar aproximadament 4.8 violacions de privacitat cada dia, amb un total de 1802 a l'any, afectant a un total de més de 422 milions de persones.

Companyies com eBay, TikTok, Reddit i corporacions com Apple, Adobe Inc. i Microsoft també han estat víctimes d'importants filtracions de dades, el que ha arribat a perjudicar a una àmplia xarxa de consumidors.

A més a més, a mesura que les dades es tornen més omnipresents en el món digital, les filtracions de dades esdevindran cada vegada més habituals. Aquest augment en el número de data breaches es pot observar en la [Figura 1.5](#), en la qual s'observa l'increment tant en el nombre de violacions de privacitat com en el nombre de persones afectades en els últims anys.

Figura 1.6: increment en el número de data breaches i persones afectades en els darrers anys.

Font: Elaboració pròpria amb dades extretes d'una de les figures de l'article: [*The effectiveness of blockchain technology in preventing financial cybercrime*](#). Publicat al maig de 2023.



GOOGLE, DON'T BE EVIL.

Google és una de les companyies més riques i influents del món que, a diferència de qualsevol altra, ha assolit una posició única de poder [Mc19] [HH13]. En efecte, Google és el mitjà amb el qual experimentem i comprenem el món que ens envolta.

Amb el propòsit d'“organitzar la informació del món i fer-la útil i accessible per a tothom” [GooSF], Google ofereix una àmplia gamma de serveis, en la seva majoria gratuïts, amb l'intenció, no només de presentar la informació d'una manera més organitzada, sinó també de recopilar la major quantitat d'informació possible [FM12]. De fet, el poder d'aquesta empresa es basa únicament en com utilitza la gran quantitat de dades que adquireix dels seus usuaris [Cas21].

GOOGLE ANDROID.

La interacció amb els servidors de Google en un dispositiu Android és 10 vegades més freqüent que la comunicació amb els servidors d'Apple des d'un dispositiu iPhone.

En termes de magnitud, els telèfons Android transmeten aproximadament 4,4 MB de dades al dia (~130 MB al mes) amb els servidors de Google.

Aquesta és una quantitat gegant, sobretot si considerem que el sistema operatiu Android és utilitzat per més de 3.000 milions d'usuaris.



Figura 1.7: Google Android i la gran recopilació de dades que efectua amb tots els dispositius que utilitzen aquest sistema operatiu.

Font: Els diversos logotips han estat extrets de la [pàgina de blog oficial de Google](#). La informació del text procedeix de l'estudi de Schemedit [Sch18].

Tanmateix, hi ha nombrosos individus que neguen o mostren indiferència davant l'ús d'informació personal que duu a terme Google i, en conseqüència, desconeixen el poder d'aquesta empresa. Malgrat la imatge neutral que Google projecta, la seva capacitat per decidir qui accedeix a quina informació i oportunitats, la situa com una entitat de considerable poder i amenaça pel que fa a la privacitat i altres aspectes.

Tot i que Google Search s'esforça per oferir resultats de cerca objectius i precisos, de vegades aquests resultats poden veure's esbiaixats o influïts pels interessos comercials de la companyia.

El 2015, la divisió Buró de Competència de la Comissió Federal de Comerç dels Estats Units (Federal Trade Commission, FTC), encarregada de l'eliminació i prevenció de pràctiques comercials anticompetitives, va publicar de manera accidental parts d'un informe sobre una investigació de Google al Wall Street Journal [Cas21]. En l'informe, la FTC va assenyalar que Google estava afavorint els seus propis productes i adoptant una estratègia anticompetitiva atès que no mostrava certes pàgines web especialitzades en categories altament comercials [Mc19] [HH13].

Nota 3: És rellevant subratllar que, malgrat el seu cèlebre lema "Don't be evil", Google continua sent una empresa comercial amb responsabilitats envers els seus accionistes i, per tant, recopila informació dels seus consumidors i l'utilitza per maximitzar els guanys [McI19].

Tot i que el seu objectiu principal és proporcionar serveis útils i accessibles, és imprescindible tenir en compte que el seu enfocament últim es dirigeix cap la generació de beneficis econòmics.

Tot seguit, en Octubre de 2020, el Subcomitè del Poder Judicial de la Cambra de Representants dels Estats Units sobre Estat Administratiu, Reforma Reguladora i Antimonopoli va publicar un informe centrat en el domini de les quatre grans empreses de dades (Google, Apple, Meta i Amazon); en el qual s'assenyalaven novament les pràctiques anticompetitives de Google³ [Cas21] [KSh20].

La informació que proporcionem a companyies tecnològiques, com Google i Meta, és extremadament personal; arribant a casos en què les companyies coneixen més atributs de la nostra pròpia personalitat que tots els nostres amics junts. En l'informe de Schmidt [Sch18], es detallen diversos experiments que evidencien l'alarmant abast de la recopilació de dades que efectua Google. La companyia recopila informació cada vegada que els usuaris interactuen amb alguna de les seves plataformes (p. ex., Chrome i Android), aplicacions (com Google Maps, YouTube i Gmail), eines per a editors (p. ex., Google Analytics i AdSense) i altres programes; cosa que resulta en un coneixement excessiu d'informació personal.

Els aspectes més confidencials i íntims de la informació emmagatzemada per Google exemplifiquen la naturalesa única de les dades com a producte de comoditat diferent de qualsevol altra mercaderia controlada per entitats centralitzades [McI19]. Les dades, a excepció de les altres mercaderies, poden ser molt valuoses per a l'elaboració de perfils massius, utilitzats per comprendre les preferències i tendències col·lectives, així com per a la creació de perfils individuals, utilitzats per interioritzar les preferències i inclinacions personals [FM12].

Si bé la recopilació i el processament de dades personals són pràctiques necessàries per millorar la qualitat i personalització dels serveis, és essencial considerar els diversos riscos de seguretat associats amb l'intercanvi d'informació sensible.

Tanmateix, malgrat les nombroses implicacions negatives, la majoria d'usuaris no només estan d'accord, sinó que inclús desitjosos de compartir les seves dades confidencials amb la finalitat de rebre serveis més personalitzats i integrats [FM12]. De fet, els consumidors estan regalant la seva informació sota un fals ideal de comoditat i cedint els seus drets de privacitat en benefici d'un servei més personalitzat i individualitzat. Això fa que els usuaris perdin la sobirania de les seves dades i quedin estrictament subjectats al poder d'aquells que les posseeixen; les companyies.



Figura 1.8: Logotip oficial de la companyia Google que s'ha utilitzat des de l'1 de setembre de 2015.

Font: La imatge ha estat extreta de la pàgina web de [Wikipedia](#).

ALGUNES DE LES APLICACIONS DE GOOGLE.

YouTube

YouTube lidera el mercat de l'streaming en línia amb un 80% de quota de mercat, allotjant més de 400 hores de nou contingut cada minut i aconseguint 1.000 milions d'hores de visualització diàries.

A mesura que els usuaris consumixen contingut, Google recopila i analitza dades detallades de l'activitat de l'usuari, proporcionant informació valuosa sobre el seu comportament.

Gmail

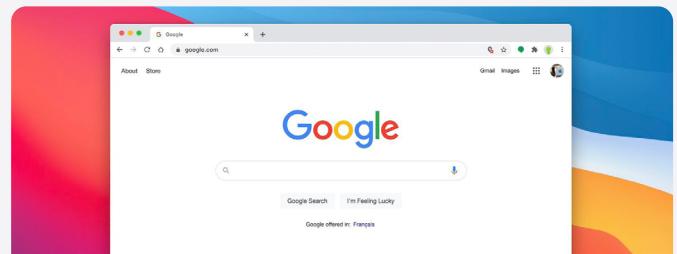
Gmail emmagatzema tots els missatges, tant enviat com rebuts, el nom del remitent, l'adreça de correu electrònic i la data/hora de tots els missatges. Com que Gmail actua com un dipòsit central de correu per a moltes persones, pot determinar els seus interessos escanejant el contingut del correu electrònic.

A més a més, ja que els usuaris poden utilitzar la seva identificació de Gmail en altres plataformes de tercers, Google pot conèixer qualsevol contingut que els provingués en forma de correu electrònic.

No obstant això, cap al final del 2017, Google va anunciar que interromprà la pràctica de personalitzar els anuncis basant-se en el contingut dels missatges de Gmail. Tanmateix, el 2018, Google va aclarir que encara escaneja els missatges de Gmail per a alguns fins.

Altres

A part d'aquestes quatre plataformes, Google compta amb un centenar més, entre les quals es troben Play Store, Google Assistant i Google Photos.



Google Search

Google Search, el motor de cerca més popular del món, gestiona més d'11.000 milions de consultes mensuals només als Estats Units.

Recopila dades de les cerques realitzades, l'historial de navegació i la interacció dels usuaris amb els anuncis. També registra la localització quan els usuaris fan servir la cerca en dispositius mòbils i d'escriptori.

Google Maps

Google Maps és la principal aplicació de navegació de Google. Pot determinar les rutes de viatge, la velocitat de moviment, el vehicle, i els llocs que un usuari freqüenta. Aquesta informació proporciona a Google informació sobre els interessos, el moviment i el comportament de l'usuari.

La precisió de la informació d'ubicació no només possibilita dirigir-se a audiències publicitàries, sinó també distribuir anuncis als usuaris quan s'apropen a les botigues.



INGENUÏTAT HUMANA.

Les tecnologies digitals han avançat de manera contínua i progressiva cap a una arquitectura centralitzada que compromet els drets dels consumidors i posa en perill la confidencialitat de la informació [FM12]. Els casos de Google i Facebook evidencien el control monopolitzat que les empreses tecnològiques exerceixen sobre la informació personal, el coneixement col·lectiu i els mitjans de comunicació. A més a més, aquests dos casos resulten ser exemples clars que il·lustren les possibles implicacions negatives de la centralització de les dades, o en altres paraules, del poder.

1. En el cas de Facebook, s'ha criticat la recopilació massiva de dades personals i el seu potencial impacte en les eleccions dels usuaris. A més a més, s'ha assenyalat que, tot i assumir que aquestes empreses no tenen altres motivacions més enllà de les econòmiques per analitzar i utilitzar les nostres dades, la seva estructura centralitzada i la seva capacitat per recopilar i emmagatzemar grans volums d'informació ha generat la possibilitat que tercers facin un ús indegit d'aquesta.
2. En referència a Google, s'ha observat amb preocupació l'extensa recopilació de dades que realitza l'empresa amb totes les seves plataformes; ja que aquestes no es limiten únicament a proporcionar informació, sinó que també tenen com a objectiu recopilar la quantitat més gran de dades possible.

Cada conjunt d'informació que les empreses són capaces d'extreure'ns les fa més poderoses. Perquè les dades, un cop processades, permeten a les companyies predir el comportament individual i col·lectiu, tot extraient grans beneficis. Aquesta gran capacitat econòmica que posseeixen les dades és la que motiva les empreses a recopilar la quantitat més gran d'informació possible sobre els seus consumidors. Com a resultat, aquesta necessitat d'informació s'ha convertit en la principal lluita pels líders tecnològics [Par11].

Per culpa de l'objectiu de les empreses de saber tant com sigui possible dels seus usuaris, no importa quan esforç faci algú per mantenir la seva informació en privat, hi ha mecanismes i eines arreu que recopilen dades personals i les comuniquen a empreses terceres, ja sigui amb el consentiment de l'individu o sense. Més sovint, però, són els mateixos usuaris els que desitgen compartir informació personal amb una àmplia varietat d'empreses interessades. La majoria dels usuaris proporcionen informació a empreses sense conèixer les implicacions de privacitat d'aquests serveis o valoren el servei i la comoditat per sobre del seu risc personal [FM12].

Per tant, és el desig de comoditat de l'ésser humà i la manca de coneixement i indiferència dels usuaris respecte a la recopilació d'informació el que dota a les empreses d'un poder inconcebible, que mai s'havia imaginat ni molt menys concentrat.

IGNORÀNCIA.

La ignorància dels ciutadans és poder pels governadors i la ignorància dels consumidors és poder per les corporacions.



Soluciones para la Privacidad

Solucions per a la Privacitat

En aquesta secció, s'exposen les solucions tant tecnològiques com jurídiques, i el seu enfoquament conjunt, per abordar el problema de la privacitat a l'era digital. A més, explorarem el Reglament General de Protecció de Dades (RGPD), la criptografia i presentarem breument un mètode criptogràfic revolucionari.



VIOLACIONS DE PRIVACITAT.

En utilitzar aplicacions en línia, els usuaris soLEN compartir una abundant quantitat d'informació personal, ja sigui en fer compres a través d'Internet o en utilitzar targetes de crèdit; en accedir a enllaços i articles, o en publicar actualitzacions personals; en qualificar sèries i pel·lícules, o en comentar en publicacions; la informació sempre es comparteix dins d'un àmbit particular. No obstant això, quan una part de la informació es mou més enllà del seu abast previst (ja sigui perquè es manté més temps de l'establert, es comparteix amb corporacions no autoritzades o s'abusa de la informació per a un propòsit diferent del pactat), es produeix una violació de la privacitat [JBE+13].

Actualment, resulta complicat conèixer amb certesa l'ús que se li està donant a les nostres dades. Però encara resulta més difícil preveure com i per qui seran usades en el futur [Stu18]. Arribant a un punt en el qual els mateixos usuaris, que són els legítims propietaris de la seva identitat, deixin de ser els veritables propietaris d'aquesta i cedeixin el poder a les empreses. L'apropiació del control de les nostres dades personals per part d'un reduït conjunt d'empreses ens priva i ens expropia de la facultat de gestionar la nostra informació confidencial; suspenent-nos del nostre dret de propietat sobre les dades que conformen la nostra identitat. En canvi, les empreses adquireixen un domini absolut o parcial sobre les nostres dades.

Així doncs, l'avancé de l'economia digital i el Big Data ha generat inquietuds entre certs usuaris que temen perdre el control sobre la manera en com es recopilen i utilitzen les seves dades. Tanmateix, altres, en prendre consciència que aquesta situació ja és una realitat, comencen a buscar alternatives per contrarestar aquestes preocupacions. En la literatura relacionada, tant empreses com investigadors han proposat diverses solucions per restituir la sobirania que es mereixen els individus. Entre aquestes solucions, les que destaquen són les jurídiques i les tecnològiques.

SOLUCIONS JURÍDIQUES.

La manca d'un marc regulador sòlid que promogui la transparència i el control del consumidor sobre les seves pròpies dades pot generar series implicacions que afectin el bon funcionament dels mercats digitals [Gia20] [Mcl19]. En resposta, s'han realitzat esforços a escala mundial per actualitzar i establir noves lleis que solucionin els reptes relacionats amb la privacitat de la informació. Exemples destacats inclouen el Reglament General de Protecció de Dades (RGPD) [Reg16] de la Unió Europea, que va entrar en vigor el 25 de maig de 2018, i la Llei de Privadesa del Consumidor de Califòrnia (CCPA), que es va implementar l'1 de gener de 2020.

La introducció del RGPD ha marcat l'inici d'una sèrie de canvis en la forma en què es gestionen les dades personals dels ciutadans europeus. Des d'una perspectiva jurídica, l'objectiu del RGPD és dotar el marc legal amb les garanties adequades que permetin el control individual sobre les dades personals [Gia20]. Aquest principi de control es reflecteix en un conjunt de drets; els més destacats són:

1. Dret de Supressió o Dret a l'Oblit: Els usuaris tenen el dret d'obtenir la supressió de les dades que els afectin. Això significa que les empreses han d'eliminar o anonimitzar les dades d'un usuari quan ja no siguin necessàries en relació amb els fins per als quals van ser recollides o quan l'interessat retiri el seu consentiment [GDa17] (Art. 17, [Reg16]).
2. Requisit del Consentiment Informat: Cada client ha de ser informat en termes senzills sobre la finalitat de les dades que proporciona [GDa17] (paràgraf 42-43, [Reg16]). A més, es requereix que els clients autoritzin prèviament l'ús de les seves dades [Mcl19] [GDa17] (Art. 6, paràgraf 1, [Reg16]).
3. Dret a la Notificació Immediata de Violacions de Privadesa: En cas de violació de la seguretat, el responsable del tractament ha d'informar als usuaris afectats de manera immediata i sense demores indegudes, com a molt tardà 72 hores després que s'hagi tingut constància de l'esdeveniment [GDa17] (Art. 33, [Reg16]).
4. Dret a la Portabilitat de les Dades: Els clients tenen dret a rebre les dades conservades sobre ells en un format estructurat, d'ús comú i lectura mecànica; i tenen dret a transmetre-les a altres organitzacions i responsables [GDa17] [Mcl19] (Art. 20, [Reg16]).

No obstant això, malgrat els avanços en aquestes legislacions, hi ha certes limitacions que limiten la seva capacitat per resoldre completament la manca de control dels usuaris.

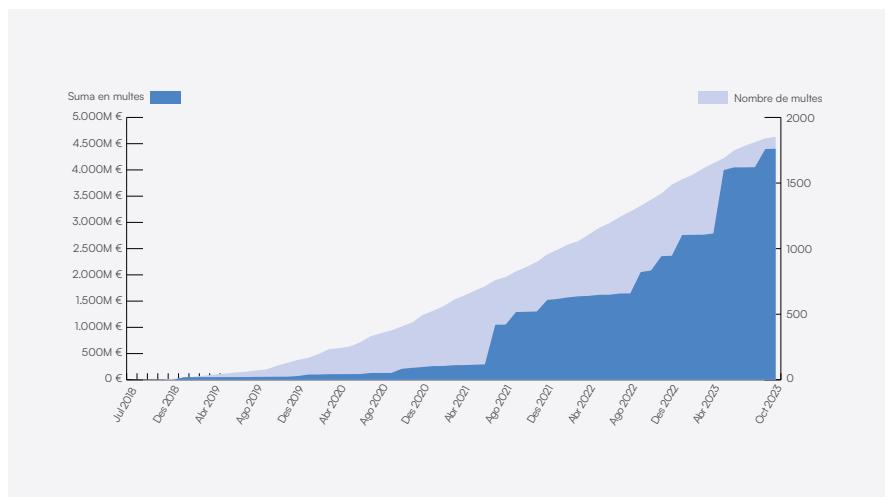


Figura 2.1: Nombre de sancions imposades pel Reglament General de Protecció de Dades (RGPD) i la suma de diners que ha estat penalitzada.

Font: Elaboració pròpia amb dades de CMS Legal Services EEIG que realitza el seguiment de l'aplicació del GDPR.

CONFIANÇA EN ABUNDÀNCIA.

El simple fet d'haver de confiar en els proveïdors de serveis o en el responsable de l'emmagatzematge planteja un desafiament en si mateix.

En l'actualitat, els usuaris depositen molta confiança implícita en els proveïdors de serveis, esperant que gestionin la seva informació de manera justa i conscient; respectant els drets del consumidor i aplicant mesures tècniques de seguretat per garantir la confidencialitat i la privadesa de les dades; i que continuïn fent-ho en el futur [JBE+13]. En utilitzar el sistema, els usuaris estableixen una relació amb el proveïdor de serveis, qui (a causa de la seva forma de funcionar) pot veure tota la informació en el sistema, incloses les descàrregues privades, les sol·licituds d'accés a informació, el comportament de compra i navegació, etc.

Per tant, correspon al mateix proveïdor assegurar-se que no s'empra aquesta informació sense el consentiment corresponent. A més, és el proveïdor de serveis qui té l'autoritat final per decidir quina informació s'emmagatzema, quant temps es conserva i com s'usa o es distribueix [JBE+13]; no és l'usuari qui decideix de forma directa.

Si bé els reglaments com el RGPD o la CCPA estableixen mesures estrictes per a aquests aspectes, és responsabilitat del proveïdor decidir si compleix o no amb aquestes mateixes normes i acceptar les conseqüències associades a la seva decisió. A més a més, poden sorgir situacions en què la responsabilitat no recaigui directament en l'ètica de l'empresa, sinó que més aviat es deu a una mala gestió o manca de recursos tècnics per abordar adequadament aquests aspectes. A tall d'exemple, la majoria de *data breaches* són causades per una gestió inadequada i no per una mala ètica; no obstant això, el resultat és el mateix: les dades que hem proporcionat amb la confiança que es mantindrien segures ja no ho estan.

Per tant, com a usuaris, no hem de depositar una confiança excessiva en les plataformes digitals. Ja que aquestes poden violar la nostra privadesa de manera voluntària en cerca de beneficis econòmics o poden ser víctimes d'una filtració de dades i exposar informació que hauria de romandre en secret. Com a mínim, no hem de concedir la nostra confiança a aquelles companyies que es basen únicament en afirmacions sense suport tecnològic, en comptes d'utilitzar plataformes dissenyades específicament per garantir una protecció completa i sense fissures.

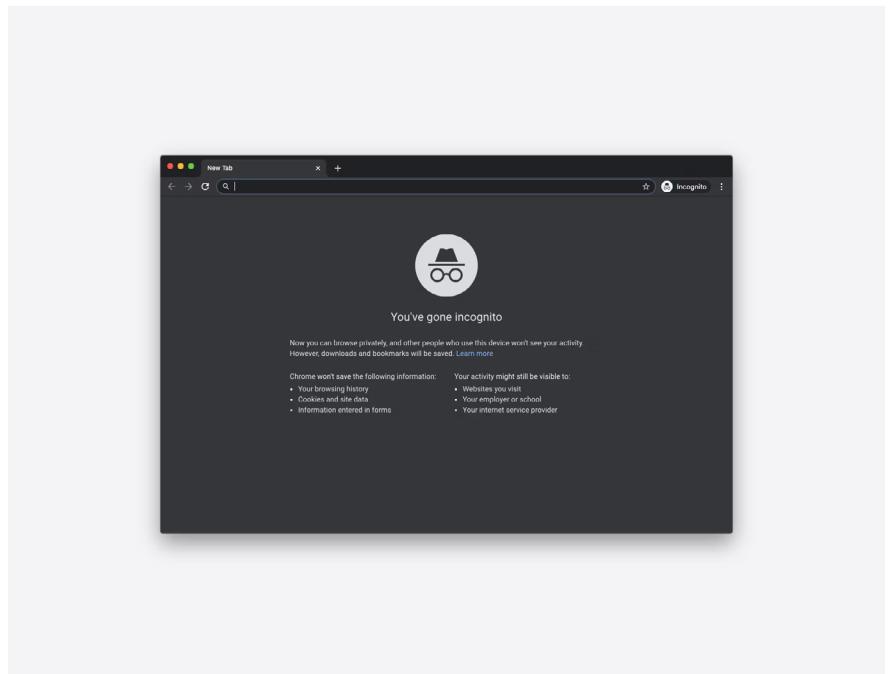


Figura 2.2: L'ús del mode incògnit impedeix que les dades o l'historial de navegació s'emmagatzemin al seu dispositiu. No obstant això, tot i que la seva activitat en línia no es guarda, els llocs web que visita, els motors de cerca i altres empreses poden continuar rastrejant el seu comportament.

Font: La imatge s'ha extret d'un dels articles penjats al blog de Setapp, concretament, l'article es titula: [How To Turn On Incognito Mode On iPhone, iPad, And Mac](#).

DECLARACIONS DE PRIVACITAT.

Generalment, les declaracions de privacitat es proporcionen per mostrar la postura adoptada pel proveïdor de serveis i obtenir el consentiment de l'usuari. No obstant això, molts usuaris es veuen forçats a acceptar els termes i condicions d'ús; perquè, en cas contrari, el servei els denegaria l'accés. Limitant el dret dels usuaris a seleccionar quina informació volen compartir i quines prefereixen mantenir privada [Stu18].

L'equilibri de poder està clarament a favor del proveïdor de serveis [JBE+13], i no es pot simplement assumir que els usuaris consenten aquesta forma de recopilació i processament de dades [Reg16]. Degut, en part, a la posició dominant del servei.

El RGPD i altres regulacions similars obligaran les empreses a detallar millor la seva política de privadesa. No obstant això, per als qui acceptin la política proposada, aquestes empreses continuaran recopilant dades de manera extensa.

Tot i que les solucions jurídiques són efectives per establir un cert nivell de privacitat, encara pateixen la debilitat inherent dels models basats en la confiança. Per això, és necessari establir un sistema que asseguri la confidencialitat de la informació a través de models basats en proves criptogràfiques en lloc de confiança.

A més, a diferència de les lleis que generalment s'empren per resoldre problemes després que sorgeixin, un enfocament basat en la tecnologia serà capaç de prevenir les violacions de privacitat abans que succeeixin [JBE+13].

LA BALANÇA DESIGUALADA.

En la gran majoria de declaracions de privacitat, no hi ha un terme mitjà. Es tracta del "toma-lo o déjalo" de les negociacions tradicionals. O acceptes l'ús que se'ls donarà a les teves dades o prescindeixes del servei.



Figura 2.3: Tot i que les declaracions de privacitat són una mesura que ofereix més opcions als usuaris, en ocasions poden jugar amb la seva posició dominant, fent que els usuaris acceptin certes declaracions que no acceptarien en altres casos.

Font: La imatge s'ha extret de Spreadshirt.

SOLUCIONS TECNOLÒGIQUES.

Les solucions tecnològiques ofereixen una major garantia de confidencialitat de la informació, ja que reemplacen la dependència de la confiança amb mecanismes matemàtics. En general, es reconeix que la tecnologia de xifratge és el mètode tècnic més simple i eficaç per protegir les dades dels usuaris [MZL+22] [FG07] [AFF+13].

Nota 4: D'acord amb el principi de Kerkhoff, la seguretat del protocol no hauria de dependre de l'opacitat del codi, sinó únicament del nivell de secret de la clau de desxifrat [FG07].

La criptografia és un camp especialitzat que se centra a protegir la confidencialitat i seguretat de la informació [MZL+22], mitjançant l'ús de tècniques de codificació i descodificació. Aquestes tècniques asseguren que la informació roman inaccessible per a aquells sense la clau de xifratge adequada⁴.

En criptografia, es distingeixen entre els esquemes de clau secreta, també coneguts com a criptografia simètrica, i els de criptografia de clau pública, també coneguts com a criptografia asimètrica. En els esquemes de criptografia simètrica, tant el xifratge com el desxifratge (d'un missatge específic, al qual anomenem m) es realitzen utilitzant la mateixa clau⁵, coneguda com a k .

$$C = Enc(k, m) \quad (1)$$

$$m = Dec(k, C) \quad (2)$$

Nota 5: Atès que s'utilitza la mateixa clau tant per xifrar com per desxifrar, és necessari mantenir-la en secret o compartir-la només amb aquelles persones amb les quals desitgem compartir informació.

En l'[Equació 1](#), la funció Enc representa la funció de xifratge, que pren una clau i un missatge sense xifrar com a paràmetres i retorna un text xifrat, conegit com a ciphertext. D'altra banda, en l'[Equació 2](#), la funció Dec representa la funció de desxifratge, que utilitza la mateixa clau que s'ha utilitzat per xifrar el missatge per tal de recuperar-ho. Com que es fa servir la mateixa clau, és necessari que l'emissor i el receptor acordin prèviament la clau que utilitzaran per establir qualsevol comunicació segura. Això implica que aquests esquemes no puguin ser emprats per dues persones que mai s'hagin conegit prèviament. A més, en aquest sistema es requereix compartir una clau diferent amb cada individu amb qui volem comunicar-nos. No obstant això, els esquemes simètrics presenten l'avantatge de ser realment ràpids i se'n fa us amb la freqüència més gran possible [FG07].

En contrast amb el model anterior, els esquemes asimètrics utilitzen un parell de claus; una de les quals, anomenada clau pública (representada com a k_p), s'utilitza per xifrar, mentre que l'altra, la clau privada (representada amb el subíndex s , k_s), es manté en secret i s'utilitza per desxifrar el ciphertext. Quan es desitja enviar un missatge xifrat, l'emissor fa ús de la clau pública del receptor per xifrar-lo. Després, el receptor farà servir la seva clau privada per desxifrar el ciphertext rebut i obtenir el missatge original.

$$C=Enc(k_p, m) \quad (3)$$

$$m=Dec(k_s, C) \quad (4)$$

L'[Equació 3](#) xifra el missatge utilitzant la clau pública i l'[Equació 4](#) el desxifra amb la clau secreta. Els esquemes asimètrics es consideren més flexibles que els simètrics, ja que no requereixen que l'emissor i el receptor acardin prèviament cap clau. Tanmateix, aquests esquemes soLEN ser més lents que els simètrics [FG07]. Exemples destacats són el RSA i l'ElGamal.

Els models criptogràfics, tant de clau pública com privada, han estat i continuen sent altament eficients per proporcionar privadesa a les plataformes. No obstant això, persisteixen certs desafiaments associats amb la criptografia emprada en l'actualitat que limiten la seva capacitat per abordar plenament la qüestió de la confidencialitat.

PROBLEMES AMB LA CRIPTOGRAFIA.

Els sistemes més coneguts d'encriptació depenen de compartir una clau, sigui pública o privada, entre els individus involucrats en l'intercanvi d'informació [AAS+18]. Ara bé, aquest enfocament planteja alguns problemes relacionats amb la privacitat.

Els usuaris o proveïdors de serveis amb accés a la clau tenen drets exclusius sobre les dades [AAS+18]. Això implica que tenen el control i la capacitat d'accedir, utilitzar i gestionar les dades personals de manera exclusiva. Especialment en el cas dels serveis en línia, existeix el risc de perdre el control sobre la confidencialitat de la informació.

Així doncs, per tal de garantir que només els usuaris legítims puguin accedir de forma segura a les seves dades, és necessari xifrar la informació abans de transmetre-la i abstendir-se de compartir qualsevol clau secreta que permeti el seu desxifratge. No obstant això, apareix un inconvenient, perquè tot i que el xifratge convencional pot ser eficaç per evitar l'accés no autoritzat, també implica la destrucció de l'estructura semàntica subjacent de les dades; cosa que fa impossible efectuar operacions sobre elles. En conseqüència, els resultats de les operacions sobre textos xifrats manquen de significat quan s'empren mètodes de xifratge tradicionals [MZL+22].

Per culpa de l'esmentat, es requereix que en qualsevol dels casos els serveis en línia hagin de desxifrar el missatge abans de fer operacions sobre ells, ja que si no, el resultat no tindrà cap significat; posant en compromís el nivell de privadesa que el servidor pot oferir als seus consumidors.

Quan es vol accedir a informació emmagatzemada en una base de dades, per exemple, l'usuari pot enviar de manera encriptada l'índex on es troba aquesta informació. No obstant això, una vegada arriba al servidor, aquest haurà de desxifrar-ho i llegir-ho per poder respondre amb la informació correcta, ja sigui amb el contingut d'una revista o amb la informació d'alguna medicina.

Nota 6: De fet, en els casos en què les intencions dels usuaris han de mantenir-se en secret, sovint els usuaris són cautelosos en accedir a bases de dades [CGE+98].

En aquest cas, existeix la possibilitat que un operador curiós segueixi les consultes de l'usuari i recol·lecti les cerques que efectua [CGE+98]; independentment del nivell de confidencialitat de la consulta realitzada o de si s'està utilitzant una pestanya d'incògnit o no⁶.

En aquest context i altres similars, els consumidors poden desitjar preservar la confidencialitat de les dades proporcionades [AFF+13]. Per exemple, un usuari pot desitjar mantenir la privacitat de la seva sol·licitud i encara així rebre el contingut sol·licitat. Això ens porta a la necessitat d'un sistema criptogràfic que respecti la confidencialitat de les dades, no només durant la comunicació i l'emmagatzematge, sinó també durant el seu processament [AFF+13].

Per descomptat, el resultat del processament ha de ser igual de bo que si les dades no estiguessin encriptades [AFF+13] [Mor13] [FG07]. De fet, el resultat desxifrat hauria de ser equivalent al resultat calculat en text sense encriptar.

Així doncs, aconseguir la coherència en els resultats de les operacions executades sobre textos xifrats representa un nou desafiament. Mentre que descobrir la solució es converteix en l'objectiu, i la resposta es transforma en el que molts criptògrafs consideren el Sant Grial de la criptografia [AFF+13] [Mor13].

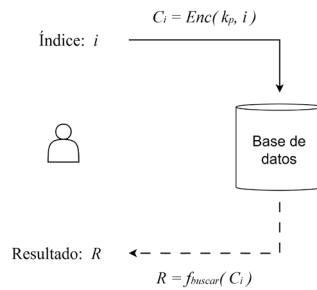


Figura 2.4: L'usuari busca accedir a la informació emmagatzemada a l'índex i. Abans d'enviar i, no obstant, xifra el seu valor utilitzant la clau pública (k_p), generant així el text xifrat de i, anomenat C_i. Un cop el servidor rep C_i, aquest el desxifra utilitzant la clau secreta (k_s) i implementa un algorisme de cerca (f_{buscar}) per trobar la informació emmagatzemada en aquest índex (R). Un cop s'obtingui, es enviarà a l'usuari. Donat que el servidor ha de desxifrar C_i per poder aplicar la funció f_{buscar}, la confidencialitat de la sol·licitud, és a dir, la privadesa de l'índex i, es perd completament.

Font: Elaboració pròpia.

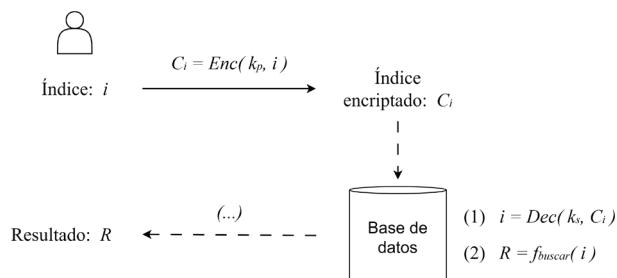


Figura 2.5: L'usuari pretén accedir a l'índex i d'una base de dades. No obstant, abans d'enviar l'índex al servidor, l'usuari l'encripta utilitzant un mètode especial de criptografia. Un cop el servidor obté el text xifrat de i (representat com a C_i), s'aplica la funció de cerca (f_{buscar}) sobre ell, sense necessitat de desxifrar-lo. Després, f_{buscar} retorna un resultat xifrat (C_R) que només l'usuari pot desxifrar, Dec(k_s, C_R), i visualitzar el seu contingut. Donat que el servidor ha pogut buscar a la base de dades l'índex i sense ni tan sols desxifrar-lo i, per tant, sense necessitat de coneixer la clau secreta de l'usuari; el servidor no ha adquirit cap informació sobre la sol·licitud ni la resposta.

Font: Elaboració pròpia.

EL SANT GRIAL DE LA CRIPTOGRAFIA.

L'any 1978, els criptògrafs Rivest, Adleman i Dertouzos van proposar una solució innovadora per abordar el desafiament de la privacitat durant el processament de dades [RAD78]. La seva proposta, coneguda avui en dia com a Xifratge Homomòrfic (Homomorphic Encryption en anglès), permetia realitzar càlculs sobre dades xifrades sense necessitat de desxifrar-les. El resultat de les operacions es retorna com un resultat xifrat, el qual, una vegada desxifrat, obté el mateix valor que si els càlculs s'haguessin efectuat sobre plaintext.

De fet, des d'aquell article seminal, el disseny de sistemes eficients i esquemes segurs d'encriptació total ha estat un dels objectius principals de la comunitat criptogràfica. Això és comprensible, ja que aquests esquemes podrien revolucionar la manera en què analitzem, manipulem i fem servir les nostres dades, oferint una forma segura i inexpugnable de processar la informació. Mitjançant aquesta tecnologia, es podria processar la informació de forma cega sense la necessitat d'accendir directament a les dades que s'estan manipulant.

Com ja vam esmentar a la Secció 1.C, els usuaris tendeixen a compartir informació personal amb una àmplia varietat d'empreses en cerca de major comoditat en el servei. Tot i que els consumidors sovint expressen la seva preocupació per la privadesa i adopten la narrativa que han de protegir-la, en examinar la realitat, es pot observar que els usuaris sovint no saben com fer-ho o no li donen tanta importància com per renunciar a algunes de les seves aplicacions favorites i més utilitzades [Hin23].

Per tant, les persones que realment poden proporcionar privadesa a tothom són els mateixos desenvolupadors d'aplicacions. Si el desenvolupador decideix implementar estrictes mesures de privadesa a la seva plataforma, qualsevol persona que l'utilitzi obtindrà seguretat i confidencialitat sense haver-se'n de preocupar. De fet, és crucial que les companyies es dediquin amb determinació a protegir la privadesa dels seus usuaris, ja que d'aquesta manera la gent no haurà de preocupar-se per la privadesa, no perquè no sigui important, sinó perquè els desenvolupadors se n'ocuparan per ells.

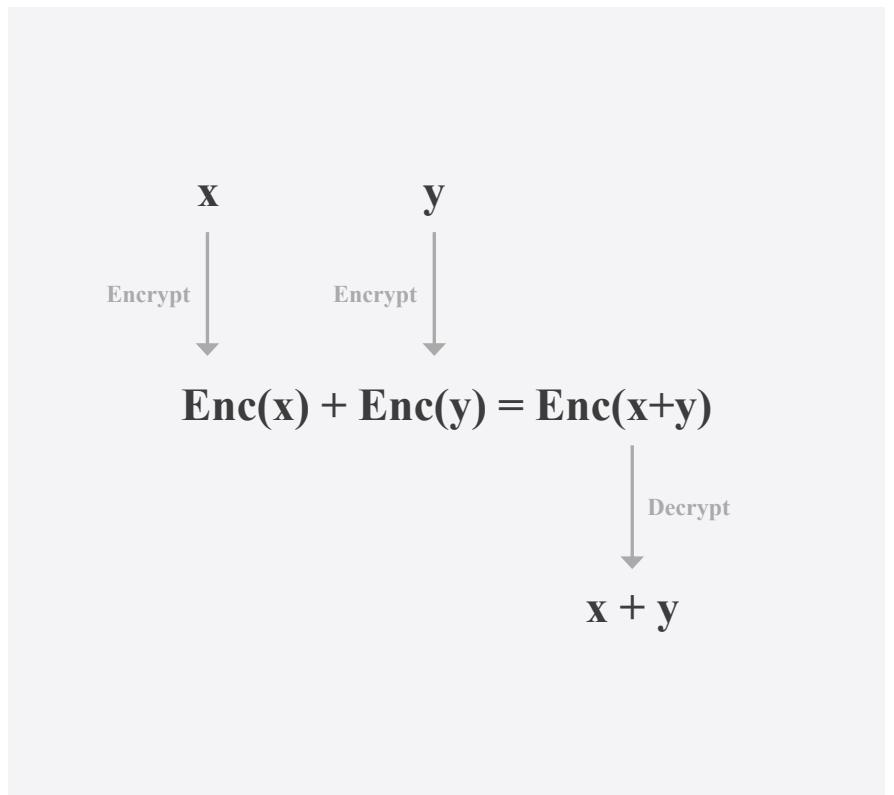


Figura 2.6: Característica de l'Homomorphic Encryption, la qual ens permet realitzar càlculs sobre dades xifrades. Dependent del sistema criptogràfic utilitzat, també es poden dur a terme operacions de multiplicació.

Font: Elaboració pròpria.

En efecte, hi ha diversos desenvolupadors i empreses que mostren una preocupació genuïna per la privadesa dels seus usuaris, ja sigui per raons ètiques o a causa de legislacions estrictes, i en alguns casos, una combinació de totes dues. No obstant això, no han aconseguit implementar models prou eficients pel simple fet que tots ells han de desxifrar el missatge abans de processar-lo.

Una empresa que analitza registres mèdics ha de conèixer el seu contingut abans de processar-los; un cercador ha de comprendre què estem cercant per oferir-nos el contingut desitjat; una plataforma de cites ha de comprendre els nostres gustos abans de començar a buscar la nostra parella ideal; els algoritmes de recomanació han de tenir coneixement de les nostres preferències per proporcionar-nos contingut que s'adapti a nosaltres, i així successivament. Totes aquestes entitats han de ser capaces d'interpretar el missatge per processar-lo i obtenir el resultat desitjat. No obstant això, això crea un punt de vulnerabilitat, ja que si les dades es mantinguessin encriptades en tot moment, incloent el processament, aquesta debilitat no existiria. Això fa que sigui difícil crear plataformes realment segures o privades.

No obstant, mitjançant l'ús de l'Encriptació Homomòrfica, és possible crear una plataforma que sigui inherentment privada des del seu disseny inicial. Això permet als desenvolupadors establir plataformes amb un nivell de seguretat del cent per cent, evitant que el servidor o l'operador en qüestió puguin accedir a cap informació sobre les nostres dades. Ja que es mantenen encriptades durant tot el procés, des que surten del nostre dispositiu fins que retorna el resultat. Brindant als consumidors la confidencialitat que desitgen sense que hagin de preocupar-se per això, ja que s'incorpora com una característica inherent al disseny de l'aplicació.

Així doncs, l'Homomorphic Encryption possibilita la implementació de sistemes totalment segurs des de la seva concepció, la qual cosa garanteix la protecció de la privadesa i la seguretat de les dades dels usuaris en totes les etapes del procés.



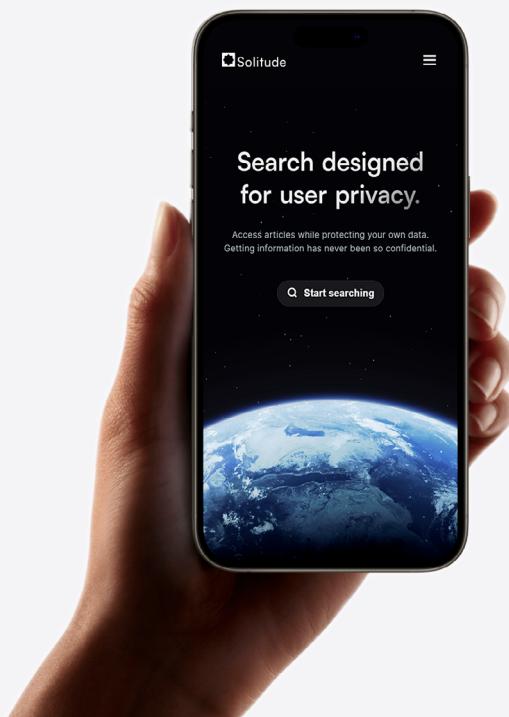
Figura 2.7: Les persones que realment poden garantir privadesa a les plataformes són les mateixes que escriuen el seu codi.

Font: Elaboració pròpria, utilitzant Adobe Photoshop.

Plataforma Creada: Solitude Search

Plataforma Creada: Solitude Search

En aquesta secció, es presentarà la part pràctica del projecte, que ha implicat el desenvolupament d'una plataforma que permet accedir a una selecció d'articles, tot assegurant la protecció de les dades dels usuaris. La plataforma en qüestió s'anomena Solitude Search.



BREU EXPLICACIÓ DE LA PLATAFORMA.

Solitude Search és un motor de cerca de bases de dades dissenyat amb l'objectiu de protegir la privadesa dels usuaris. Mitjançant aquesta eina, qualsevol persona pot accedir a una àmplia varietat d'articles de manera confidencial. Totes les sol·licituds d'accés a informació es xifren de manera que ni el servidor, ni cap altra entitat externa, té accés a les cerques dels consumidors.

Amb la finalitat contrarestar la falta de privacitat en Internet, Solitude Search es recolza en els principis de l'*Homomorphic Encryption* per garantir la Recuperació d'Informació Privada (*Private Information Retrieval*, PIR). És a dir, implementa algoritmes homomòrfics que asseguren que els usuaris aconsegueixin la informació que cerquen sense revelar cap detall personal o específic de la seva consulta.

Si bé en l'actualitat la majoria dels servidors utilitzen protocols criptogràfics per protegir els missatges durant la seva transmissió, la protecció durant el processament continua sent un repte complicat que està en continua evolució. Malgrat tot, Solitude Search es destaca com una de les pioneres en la implementació d'algorismes criptogràfics que permeten processar les sol·licituds d'informació de manera xifrada, gràcies al suport del codi obert proporcionat per l'empresa Blyss.

Encara que el seu rendiment és lent en comparació amb els competidors que no implementen aquest tipus de protocols, la plataforma compleix el seu propòsit a la perfecció, ja que estableix un sistema privat i segur des del seu disseny inicial.

És ben sabut que grans plataformes com eBay, TikTok, Facebook, Reddit i corporacions com Apple, Adobe Inc. i Microsoft han estat afectades per importants filtracions de dades. Tanmateix, Solitude Search esdevé completament segur enfront aquestes

amenaces, ja que les dades romanen xifrades durant tot el procés i ningú, excepte el mateix usuari, coneix el contingut de la informació.

Cal remarcar també que els models criptogràfics utilitzats per Spiral SDK, el kit de desenvolupament implementat per aconseguir el xifratge homomòrfic, estan basats en reticles (lattice-based cryptography) [MW22]. Els quals estan sent investigats pel National Institute of Standards and Technology (NIST) com una de les criptografies prominents que podrien resistir als atacs d'ordinadors quàntics [Bar21]. En conseqüència, la plataforma també adquireix aquesta característica, fent que les sol·licituds dels usuaris esdevinguin resistentes a les computadores quàntiques.

Així doncs, mitjançant proves criptogràfiques, la pàgina web dissenyada per aquest Treball de Recerca estableix un model de cerca completament confidencial i segur que restitueix la sobirania de les dades per part dels usuaris.

EINES EMPRADES

Per a la creació de Solitude Search, s'han utilitzat diversos llenguatges de programació amb característiques i metodologies diferents com Rust, JavaScript, HTML, CSS i Python, junt amb projectes de codi obert com Actix Web, NGINX i Spiral Privacy Clients (de l'empresa Blyss):

Nota 7: Un llenguatge de programació amb característiques de baix nivell és aquell en què les seves instruccions exerceixen un control directe sobre l'hardware.

1. **Rust:** És un llenguatge de baix nivell⁷ que es caracteritza per la seva seguretat i el seu alt rendiment. S'ha utilitzat amb Actix Web per a la creació del servidor, en el qual s'ha incorporat el codi de Spiral Privacy Clients, desenvolupat en Rust, per a la implementació d'un sistema PIR. A més, per xifrar les consultes amb un model d'*Homomorphic Encryption* abans de transmetre-les per HTTP o HTTPS, s'ha processat parts del codi de Rust amb WebAssembly per a la seva execució en el navegador.
2. **WebAssembly (Wasm):** Tot i que fins ara el desenvolupament web estava limitat a l'ús de CSS, HTML i JavaScript, la incorporació de WebAssembly ha permès l'execució de codi escrit en altres llenguatges com C, C++ i Rust en els mateixos navegadors web. En aquest cas, s'ha fet servir per a executar codi programat amb Rust que permet el xifratge del missatge directament en el navegador. (Wasm no es considera un llenguatge de programació ni un projecte de codi obert, sinó una eina per al desenvolupament web.)
3. **Spiral Privacy Clients:** És un projecte sota la llicència MIT dissenyat per l'empresa Blyss, la qual està especialitzada en la Recuperació d'Informació Privada mitjançant algorismes d'*Homomorphic Encryption*. El repositori ha permès incorporar un sistema de xifratge homomòrfic i de processament de sol·licituds en la plataforma, per tal de garantir la privacitat de les cerques dels usuaris.
4. **Actix Web:** És un *framework*⁸ per aplicacions web desenvolupat en el llenguatge de programació Rust. Està dissenyat per crear aplicacions web d'alt rendiment i escalabilitat. És conegut per ser un dels *framework* web més ràpids del món. A més, com Spiral Privacy Clients està implementat en Rust, facilita la incorporació del programari en la plataforma. No obstant això, el servidor d'Actix s'ha implementat amb HTTP, la versió no segura del protocol de transferència d'hipertext. Així que per assegurar una connexió segura entre els clients i el servidor, s'ha hagut d'implementar una *reverse proxy*⁹ amb NGINX per habilitar HTTPS, la versió segura d'aquest protocol.
5. **NGINX:** Actualment, un dels servidors web i *reverse proxy* més populars és NGINX, que està dissenyat per gestionar el tràfic web, millorar el rendiment i oferir més seguretat a les pàgines. En aquest cas, NGINX és l'encarregat de lliurar els continguts HTML, CSS i JavaScript (entre altres dades) al navegador del client, actuant com un servidor convencional. No obstant això, quan es realitza una cerca, el servidor NGINX es comporta com un servidor proxy invers (*reverse proxy*),

Nota 8: Un *framework* és un conjunt de regles i convencions que s'utilitzen per desenvolupar programari de manera més eficient i ràpida.

Nota 9: Una *reverse proxy* és un servidor intermediari que redirigeix les sol·licituds dels clients cap a servidors web interns. Actua com a capa de protecció i gestió del trànsit, millorant la seguretat i la velocitat dels servidors.

redirigint les sol·licituds del client al servidor d'Actix intern que s'executa en local (localhost), on es processen les cerques i es retorna la informació pertinent. A més, perquè NGINX funcioni amb HTTPS, s'ha adquirit un certificat SSL i s'ha implementat al servidor, permetent així que la web funcioni amb la versió segura de HTTP.

6. HyperText Markup Language (HTML): Un dels arxius que el servidor proporciona al navegador del client té l'extensió HTML. Aquest arxiu s'utilitza per crear i estructurar pàgines web. Amb l'ús d'etiquetes, s'estableix l'estructura del contingut, que pot incloure text, imatges, enllaços i elements multimèdia. També facilita la creació de pàgines interactives amb l'ajut de JavaScript i permet la personalització de l'estètica a través de CSS.
7. Cascading Style Sheets (CSS): És un llenguatge de programació usat per a definir la presentació i l'estil de les diverses pàgines. A través de regles i propietats, permet controlar l'aparença del contingut, incloent-hi el disseny, els colors, les mides de text, l'espaiat i molts altres aspectes visuals.
8. JavaScript (JS): A diferència de CSS i HTML, JavaScript és emprat per a recuperar i enviar informació al servidor, així com per a totes les operacions relacionades amb el tractament de la informació i l'aplicació d'algorismes en l'ordinador del client. També s'ha fet ús d'ell per a injectar elements HTML a la pàgina web i per a totes les funcions relacionades amb la interactivitat de l'usuari.
9. Python: S'ha aplicat exclusivament per comprimir i agrupar tots els articles en paquets de la mateixa grandària, amb la finalitat de preparar-los per ser processats per un script de Rust. Aquest és un procés que es realitza fora del servidor.

Tot i que aquests són els elements principals de Solitude Search, cadascun té matisos i utilitats addicionals que permeten que la plataforma funcioni.

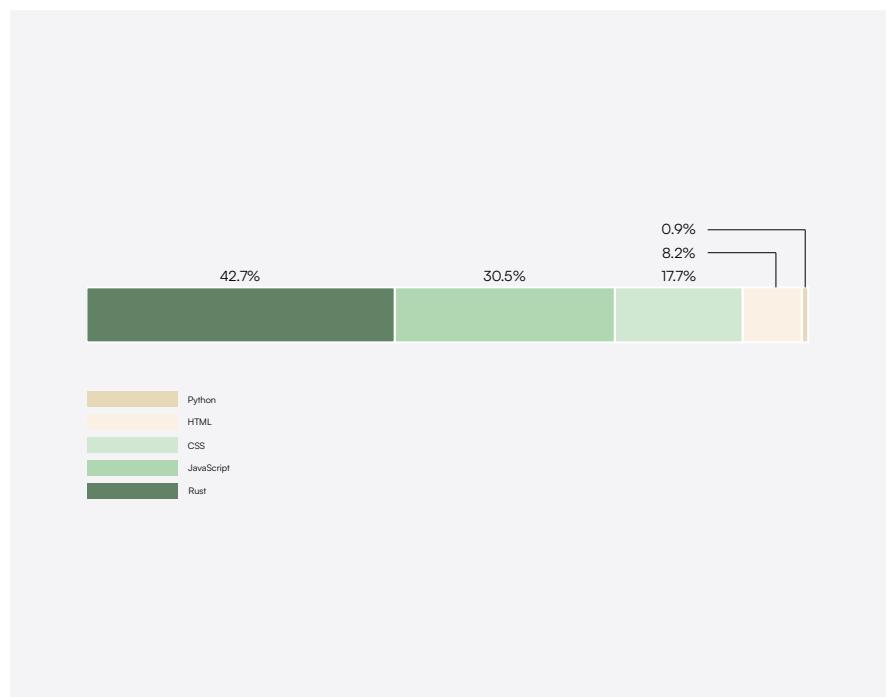


Figura 3.1: Gràfic amb els percentatges dels llenguatges de programació més utilitzats a l'hora de crear Solitude Search.

Font: Elaboració pròpia.

PROGRAMACIÓ.

El codi de Solitude Search consisteix en més de 17.000 línies de programació.

Tenint en compte que una pàgina usualment conté 40 línies en Arial 12, el codi equivaldría a unes 425 pàgines. A raó de no saturar el treball, tot el codi del programa, així com les seves carpetes, imatges i tipografies, es troben disponibles a GitHub sota la llicència MIT. L'enllaç es el següent:

github.com/Gasofa06/Solitude

A continuació, només es presenta el codi del servidor Actix com a exemple, el qual està programat en Rust.

```
1 use futures::StreamExt;
2 use spiral_rs::aligned_memory::*;
3 use spiral_rs::client::*;
4 use spiral_rs::params::model_parameters::*;
5 use spiral_rs::server::*;

6
7 use spiral_rs::util::{ params_from_json };
8 use spiral_rs::params::project_parameters::{ PROJECT_PARAMETERS };

9
10 use std::collections::HashMap;
11 use std::collections::VecDeque;
12 use std::env;
13 use std::fs::File;
14 use std::sync::Mutex;
15
16 use actix_cors::Cors;
17 use actix_web::error::PayloadError;
18 use actix_files::{ Files, NamedFile };
19 use actix_web::{ get, http, middleware, post, web, App, HttpServer };
20 use serde::Deserialize;
21
22 const PUB_PARAMS_MAX: usize = 250;
23
24 struct ServerState<'a> {
25     fname: String,
26     params: &'a Params,
27     db: AlignedMemory64,
28     pub_params_map: Mutex<
29         (VecDeque<String>, HashMap<String, PublicParameters<'a>>)
30     >,
31 }
```

```

32
33     async fn get_request_bytes(
34         mut body: web::Payload,
35         sz_bytes: usize
36     ) -> Result<Vec<u8>, http::Error> {
37         let mut bytes = web::BytesMut::new();
38         while let Some(item) = body.next().await {
39             let item_ref = &item?;
40             bytes.extend_from_slice(item_ref);
41
42             if bytes.len() > sz_bytes {
43                 println!(“too big! {}”, sz_bytes);
44                 return Err(PayloadError::Overflow.into());
45             }
46         }
47         Ok(bytes.to_vec())
48     }
49
50     fn get_other_io_err() -> PayloadError {
51         PayloadError::Io(std::io::Error::from(std::io::ErrorKind::Other))
52     }
53
54     fn other_io_err<T>(_: T) -> PayloadError {
55         get_other_io_err()
56     }
57
58     fn get_not_found_err() -> PayloadError {
59         PayloadError::Io(std::io::Error::from(std::io::ErrorKind::NotFound))
60     }
61
62     #[derive(Deserialize)]
63     pub struct CheckUuid {
64         uuid: String,
65     }
66
67     #[get(“/api/check”)]
68     async fn check<’a>(
69         web::Query(query_params): web::Query<CheckUuid>,
70         data: web::Data<ServerState<’a>>
71     ) -> Result<String, http::Error> {
72         let pub_params_map = data.pub_params_map.lock().map_err(other_io_err)?;
73         let has_uuid = pub_params_map.1.contains_key(&query_params.uuid);
74         Ok(
75             format!(“{\\“uuid\\”: “{}\\”, “is_valid\\”: {}}”, query_params.uuid, has_uuid)
76         )
77     }
78
79     #[post(“/api/setup”)]
80     async fn setup<’a>(
81         body: web::Bytes,
82         data: web::Data<ServerState<’a>>
83     ) -> Result<String, http::Error> {
84         println!(“Seting up some query...”);
85
86         let pub_params = PublicParameters::deserialize(data.params, &body);
87
88         let uuid = uuid::Uuid::new_v4();
89         let mut pub_params_map = data.pub_params_map.lock().map_err(other_io_err)?;
90         pub_params_map.0.push_back(uuid.to_string());
91
92         println!(“Hi”);

```

```

93     pub_params_map.l.insert(uuid.to_string(), pub_params);
94
95     if pub_params_map.l.len() > PUB_PARAMS_MAX {
96         let lru_uuid_str = pub_params_map.0.pop_front().ok_or(get_other_io_err())?;
97         pub_params_map.l.remove(&lru_uuid_str);
98     }
99 }
100
101 Ok(format!("{{\"id\":\"{}\"}}", uuid.to_string()))
102 }
103
104 const UUID_V4_STR_BYTES: usize = 36;
105
106 #[post("/api/query")]
107 async fn query<'a>(
108     body: web::Payload,
109     data: web::Data<ServerState<'a>>
110 ) -> Result<Vec<u8>, http::Error> {
111     let request_bytes = get_request_bytes(
112         body,
113         UUID_V4_STR_BYTES + data.params.query_bytes()
114     ).await?;
115     let uuid_bytes = &request_bytes.as_slice()[..UUID_V4_STR_BYTES];
116     let data_bytes = &request_bytes.as_slice()[UUID_V4_STR_BYTES..];
117     let uuid = uuid::Uuid
118         ::try_parse_ascii(uuid_bytes)
119         .map_err(|_| PayloadError::EncodingCorrupted)?;
120
121     let pub_params_map = data.pub_params_map.lock().map_err(other_io_err)?;
122     let pub_params = pub_params_map.l
123         .get(&uuid.to_string())
124         .ok_or(get_not_found_err())?;
125
126     let query = Query::deserialize(data.params, data_bytes);
127
128     let result = process_query(
129         data.params,
130         pub_params,
131         &query,
132         data.db.as_slice()
133     );
134
135     Ok(result)
136 }
137
138 #[actix_web::main]
139 async fn main() -> std::io::Result<()> {
140     println!("Starting server.");
141
142     let args: Vec<String> = env::args().collect();
143     let db_preprocessed_path = &args[3];
144     let port = &args[4];
145
146     println!("The database preprocessed path is '{}', db_preprocessed_path);
147     println!("The server will be listening on port {}.", port);
148
149     let cfg_expand = &PROJECT_PARAMETERS;
150     println!("{}", cfg_expand);
151     let box_params = Box::new(params_from_json(cfg_expand));
152     let params: &'static Params = Box::leak(box_params);
153

```

```

154     println!("\\nLoading preprocessed database...");  

155     let mut file = File::open(db_preprocessed_path).unwrap();  

156     let db = load_preprocessed_db_from_file(params, &mut file);  

157     println!("Done loading.");  

158  

159     let server_state = ServerState {  

160         fname: db_preprocessed_path.clone(),  

161         params: params,  

162         db: db,  

163         pub_params_map: Mutex::new((VecDeque::new(), HashMap::new())),  

164     };  

165  

166     let state = web::Data::new(server_state);  

167  

168  

169     let cors_fn = || {  

170         Cors::default()  

171             .allow_any_origin()  

172             .allowed_headers([  

173                 http::header::ORIGIN,  

174                 http::header::CONTENT_TYPE,  

175                 http::header::ACCEPT,  

176             ])  

177             .allowed_methods(vec!["GET", "POST", "OPTIONS"])  

178             .max_age(3600)  

179     };  

180  

181     let app_build = move ||  

182         App::new()  

183             .app_data(state.clone())  

184             .app_data(web::PayloadConfig::new(1 << 32))  

185             .service(setup) // POST  

186             .service(query) // POST  

187             .service(check) // GET  

188  

189     HttpServer::new(app_build)  

190         .bind(("localhost", port.parse().unwrap()))  

191         .unwrap()  

192         .run().await  

193 }

```

Val la pena recordar que aquest és només el codi del servidor intern; posteriorment, hi ha el servidor NGINX i molts altres elements que omplirien un bon nombre de pàgines. Per veure tots els fitxers, si us plau, dirigu-vos a [GitHub](#).

DISSENY DE LA PÀGINA INICIAL.

Solitude Search

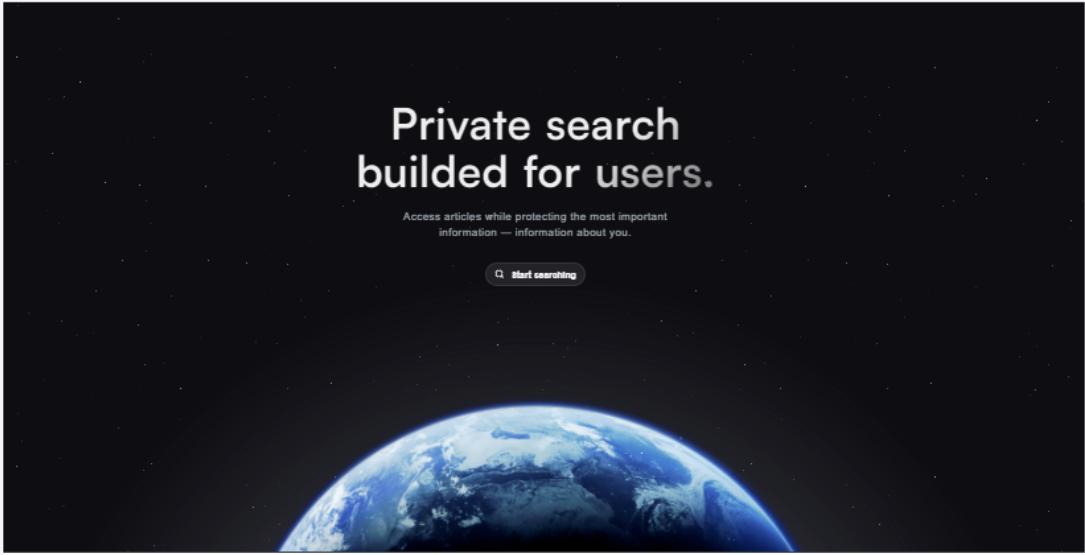
Search About FAQ

Login Back

Private search builded for users.

Access articles while protecting the most important information — information about you.

Start searching



REAGUTSAR! Project Description Benefits Search Types FAQs Acknowledgments

BRIEF ABOUT

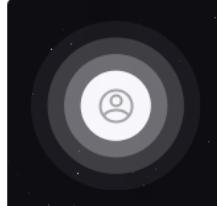
Solitude Search

Secure Retrieval Confidential Query Data Sovereignty

Incognito dB is a database search engine designed to ensure user privacy. With it, you can access hundreds of articles with an unparalleled level of confidentiality. The requests you make will be encrypted in such a way that neither the server itself nor a third-party company will be able to learn about your browsing behavior.

Designed with Spiral and based on SpiralWiki, Incognito dB is a platform that aims to counteract the abuse of data collection by companies through technology.

Browse Articles



Power must move back from the corporation to the individual.

We are expected to be the owners of our personal data, however, in order to minimally participate in modern life, we are forced to renounce this fundamental right.

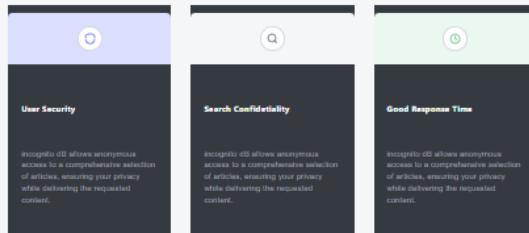
Empower Users Read the Docs →

PROJECT DESCRIPTION

Why and how

Incognito dB didn't just appear overnight; rather, it's a response to the alarming trend on the Internet of treating user information as mere raw material for exploitation. The value of Incognito dB, along with its supporting technologies, lies in safeguarding user privacy and sovereignty.

Benefits.

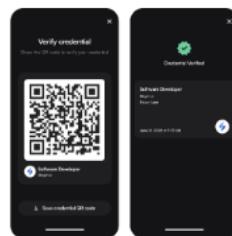


Select your preferred type of search.

Search With Title

Cloudflare IPFS Gateway provides a bridge between the current Web 2.0 model and the new decentralized and trustless model of Web3. With Cloudflare IPFS Gateway, customers can start leveraging IPFS and make content on IPFS easily accessible via HTTP. Customers can use Cloudflare DNS to map IPFS content to a domain name.

[Search by title](#)



Search With Topics

Cloudflare IPFS Gateway provides a bridge between the current Web 2.0 model and the new decentralized and trustless model of Web3. With Cloudflare IPFS Gateway, customers can start leveraging IPFS and make content on IPFS easily accessible via HTTP. Customers can use Cloudflare DNS to map IPFS content to a domain name.

[Search by topic](#)

Any Questions? Look Here

What's new?

What's new?

What's new?

DISSENY DEL CERCADOR.

Solitude Search

Search About Contact Login Dark

Solitude Search

TITLE ▾ | Search article by title (e.g. Conan Edogawa) Articles loaded.

Solitude Search

Search About Contact Login Dark

← Return to Homepage

Centralización del Poder

By Roger Rovira & Peter Parker & Al Hollings November 2023

Summary
Los medios de círculos centralizados siguen siendo predominantes en el mercado, debido a la prevalencia de los servicios pensados en Internet [72]. El propósito de esta sección es informar al lector acerca de la inequidad acumulativa de poder que emerge de la centralización de la información.

In This Article
Header
Summary
Era Digital
Datos Personales
Propaganda de Información
Género
Facebook y Cambridge Analytica

I. Era Digital
En las últimas dos décadas, la humanidad ha sido testigo del rápido desarrollo de tecnologías digitales capaces de mejorar el bienestar humano, así como ha presenciado una creciente concentración de dichas tecnologías en manos de un reducido grupo de potencias tecnológicas, como Google, Apple, Meta, Amazon y Microsoft, comúnmente conocidas como las Big Tech [14] [25] [37].

Daily Facebook Users

A donut chart titled "Daily Facebook Users" showing the percentage of global population that uses Facebook daily. The chart is divided into three segments: Global Population (grey), Facebook Users (blue), and Daily Facebook Users (yellow). The yellow segment represents 25% of the total population.

| | Google | Bing | Yahoo! |
|------|--------|------|--------|
| 2010 | 90.91 | 3.48 | 3.93 |
| 2014 | 89.81 | 3.83 | 3.57 |
| 2018 | 91.4 | 2.82 | 2.15 |
| 2022 | 92.07 | 3.19 | 1.38 |

Hace veinticinco años Meta no existía, Google era apenas un proyecto universitario y Amazon solo se enfocaba en la venta de libros [4], y, sin embargo, son actualmente consideradas como algunas de las empresas más valiosas e influyentes del mundo. Dada su inmensa magnitud y alcance, estas compañías tienen la capacidad de influir sobre el comportamiento de los usuarios y la economía internacional [37], lo que les otorga un poder sin precedente que se manifiesta en varios aspectos de nuestra vida cotidiana.

Un ejemplo relevante es el caso de Cambridge Analytica, en el cual se puso de manifiesto la función que desempeñaron los datos de Facebook en las quincuagésimas octavas votaciones presidenciales de los Estados Unidos [46], demostrando así la gran repercusión que las grandes empresas tecnológicas ejercen en la opinión pública y los procesos democráticos.

Del mismo modo que Meta se ha apoderado de la información de sus usuarios, tanto Google como Twitter han sido capaces de monopolizar la información de diversas maneras [44], colocándose en una posición privilegiada para capitalizar la innovación futura y ofrecer servicios personalizados.

Conclusions Finals del Projecte

Conclusions Finals del Projecte



CONCLUSIÓ.

És probable que l'arribada d'Internet i la tecnologia digital sigui la característica definitòria d'aquest segle. La seva promesa és colossal com ho és el seu potencial d'ús indegit. Si les dades han de ser el motor d'aquesta nova tecnologia, les hem de gestionar bé amb la finalitat d'evitar la seva mala utilització.

No obstant la importància que representa, la seguretat de les dades personals ha estat un problema constant en les tecnologies de la informació. Si bé s'han implementat diverses solucions, cap d'elles ha estat capaç de dotar als usuaris amb les garanties necessàries que permetin el control sobre les seves dades i atorguin privacitat al sistema. És més, a causa de la lluita entre empreses d'obtenir quelcom més d'informació, la privacitat en Internet no ha fet més que disminuir. En aquest sentit, el concepte d'*Homomorphic Encryption* ha guanyat notorietat, ja que és l'única tècnica criptogràfica que ofereix el mateix nivell de privadesa que altres criptografies, alhora que permet efectuar operacions sobre les dades sense haver de desxifrar-les. Així és que pot oferir privadesa en l'era de la informació.

Per demostrar la seva viabilitat i proporcionar una solució específica al problema de la privacitat, s'ha dissenyat una plataforma de Recuperació d'Informació Privada. A través d'aquesta, els clients poden accedir a articles emmagatzemats en una base de dades sense la necessitat de comprometre cap informació sobre la seva cerca. Malgrat que la plataforma incorpora un cercador de bases de dades totalment confidencial, es pot notar un increment en els temps de resposta del servidor, ja que és necessari realitzar un major nombre de càlculs i de major complexitat per mantenir la informació de l'usuari en secret. El camp de l'*Homomorphic Encryption* i la Recuperació d'Informació Privada està en constant evolució i és probable que en un futur els algoritmes millorin, assolint així una reducció en els càlculs necessaris.

Tanmateix, aquest treball ha presentat una de les primeres plataformes en implementar un sistema totalment privat. Encara que el seu rendiment és lent en comparació amb els competidors, cap altre ofereix el mateix nivell de confidencialitat.

És hora de ser els propietaris de les nostres dades i Solitude Search és un mitjà per aconseguir-ho.

Referències

Referències

A continuació es mostren totes les referències utilitzades en la redacció del text. Les referències d'imatges que s'han extret de certes fonts i no s'han creat expressament per a aquest projecte normalment tenen la cita just al seu costat.



ARTICLES.

- [AFF+13]** Aguilar-Melchor, C., & Fau, C., & Fontaine, C., & Gogniat, G., & Sirdey, R. (2013, febrer 13). Recent Advances in Homomorphic Encryption: A Possible Future for Signal Processing in the Encrypted Domain. Recuperat de <https://ieeexplore.ieee.org/document/6461628>.
- [AAS+18]** Acar, A., & Aksu, H., & Selcuk, A., & Conti, M. (2018, juliol 25). A Survey on Homomorphic Encryption Schemes: Theory and Implementation. Recuperat de <https://dl.acm.org/doi/10.1145/3214303>.
- [Bar21]** Barak, B. (2021, novembre 17). An Intensive Introduction to Cryptography: [Chapter 11] Lattice based cryptography. Recuperat de https://intensecrypto.org/public/lec_12_lattices.html (Consultat el 20 d'octubre de 2023).
- [BC21]** Birch, K., & Cochrane, D. T. (2021, maig 26). Big Tech: Four Emerging Forms of Digital Rentiership. Recuperat de <https://doi.org/10.1080/09505431.2021.1932794>.
- [Cam05]** Cameron, K. (2005, maig 11). The Laws of Identity. Recuperat de <https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf> (Consultat el 26 de març de 2023).
- [Cas21]** Case, M. (2021, agost 15). Google, Big Data, & Antitrust. Recuperat de <https://ssrn.com/abstract=3917218>.
- [CGE+98]** Chor, B., & Goldreich, O., & Eyal, K., & Sudan, M. (1998, novembre 1). Private information retrieval. Recuperat de <https://dl.acm.org/doi/10.1145/293347.293350>.
- [FG07]** Fontaine, C., & Galand, F. (2007, octubre 24). A Survey of Homomorphic Encryption for Nonspecialists. Recuperat de <https://doi.org/10.1155/2007/13801>.
- [FM12]** De Filippi, P., & McCarthy, S. (2012, octubre 26). Cloud Computing: Centralization and Data Sovereignty. Recuperat de <https://ssrn.com/abstract=2167372>.
- [GDa17]** G Data Software AG. (2017, setembre). El nuevo Reglamento de Protección de Datos de la UE (GDPR) – Lo que las empresas deben saber. Recuperat de https://file.gdatasoftware.com/web/es/documents/whitepaper/G_DATA_El_nuevo_Reglamento_de_Proteccion_de_Datos_de_la UE_GDPR.pdf (Consultat el 28 de juny de 2023).

- [Gia20]** Giannopoulou, A. (2020, setembre 28). Data Protection Compliance Challenges for Self-Sovereign Identity. Recuperat de <https://dx.doi.org/10.2139/ssrn.3671523>.
- [HH13]** Haucap, J., & Heimeshoff, U. (2013, agost 21). Google, Facebook, Amazon, eBay: Is the Internet driving competition or market monopolization? Recuperat de <https://doi.org/10.1007/s10368-013-0247-6>.
- [JBE+13]** Jeckmans, A., & Beye, M., & Erkin, Z., & Hartel, P., & Lagendijk, R., & Tang, Q. (2013). Privacy in Recommender Systems. Recuperat de https://ris.utwente.nl/ws/portalfiles/portal/5352108/Privacy_in_Recommender_Systems.pdf (Consultat el 24 de juny de 2023).
- [McI19]** McIntosh, D. (2019, gener). We Need to Talk About Data: How Digital Monopolies Arise and Why They Have Power and Influence. Recuperat de <https://scholarship.law.ufl.edu/jtlp/vol23/iss2/2>.
- [Mor13]** Morris, L. (2013, maig 10). Analysis of Partially and Fully Homomorphic Encryption. Recuperat de <http://gauss.ececs.uc.edu/Courses/c5156/pdf/homo-outline.pdf> (Consultat el 9 de juliol de 2023).
- [MTM07]** Mossberger, K., & Tolbert, C. J., & McNeal, R. S. (2007). Digital Citizenship: The Internet, Society, and Participation. Recuperat de <https://doi.org/10.7551/mitpress/7428.001.0001>.
- [MW22]** Menon, S. J., & Wu, D. J. (2022). Spiral: Fast, High-Rate Single-Server PIR via FHE Composition. Recuperat de <https://ia.cr/2022>.
- [MZL+22]** Ma, Y., & Zhao, J., & Li, K., & Cao, Y., & Chen, H., & Zhang, Y. (2022, octubre 4). Research Review on the Application of Homomorphic Encryption in Database Privacy Protection. Recuperat de <https://doi.org/10.4018/IJCINI.287600>.
- [RAD78]** Rivest, R., & Adleman, L., & Dertouzos, M. (1978). On data banks and privacy homomorphisms. Recuperat de <https://cdn.sanity.io/files/r000fwn3/production/c365f01d330b2211e74069120e88cff37eacbcf5.pdf> (Consultat el 9 de juliol de 2023).
- [Sch18]** Schmidt, D. C. (2018, agost 15). Google Data Collection Research. Recuperat de <https://digitalcontentnext.org/blog/2018/08/21/google-data-collection-research> (Consultat el 20 de maig de 2023).
- [SG17]** Stucke, M. E., & Grunes, A. P. (2017, març 3). Data-opolies. Recuperat de: <https://dx.doi.org/10.2139/ssrn.2927018>.
- [Shr20]** Shrestha, A. K., & Vassileva, J., & Deters, R. (2020, octubre 22). A Blockchain Platform for User Data Sharing Ensuring User Control and Incentives. Recuperat de <https://doi.org/10.3389/fbloc.2020.497985>.
- [Stu18]** Stucke, M. E. (2018, març 19). Should We Be Concerned About Data-opolies? Recuperat de <https://dx.doi.org/10.2139/ssrn.3144045>.
- [Yoo12]** Yoo, C. S. (2012, gener 7). When Antitrust Met Facebook. Recuperat de https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1421&context=faculty_scholarship.
- [ZTL+18]** Zhan, Y., & Tan, K. H., & Li, Y., & Tse, Y. K. (2018, novembre). Unlocking the power of big data in new product development. Recuperat de <https://doi.org/10.1007/s10479-016-2379-x>.

DOCUMENTS LEGALS.

- [Alp22]** Alphabet Investor Relations. (2022, febrer 1). FORM 10-K. Recuperat de <https://abc.xyz/assets/d9/85/b7649a9f48c4960ad bce5bd9fb54/20220202-alphabet-10k.pdf>.
- [Met23]** Meta Investor Relations (2023, febrer 2). FORM 10-K. Recuperat de <https://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/e574646c-c642-42d9-9229-3892b13aabfb.pdf>.
- [Reg16]** Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). Recuperat de <https://www.boe.es/DOUE/2016/119/L00001-00088.pdf>.

PÀGINES WEB.

- [Goo22]** Google – Privacy & Terms. (2022, desembre 15). Google Privacy Policy. Recuperat de <https://policies.google.com/privacy> (Consultat el 29 d'abril de 2023).
- [GooSF]** Google (s.f.). Sobre Nosotros. Recuperat de <https://about.google/>.

ARTICLES PERIODÍSTICS EN LÍNIA.

- [Aie18]** Aiello, C. (2018, abril 4). Cambridge Analytica says no more than 30 million people impacted by leak. CNBC. Recuperat de <https://www.cnbc.com/2018/04/04/cambridge-analytica-says-no-more-than-30-million-people-impacted.html> (Consultat el 20 de maig de 2023).
- [Cad18]** Cadwalladr, C., & Graham-Harrison E. (2018, març 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian. Recuperat de <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>.
- [IBM20]** IBM Research. (2020, desembre 16). 5 Things to Know About IBM's New Tape Storage World Record. IBM Newsroom. Recuperat de <https://newsroom.ibm.com/IBM-research?item=32682#> (Consultat el 8 d'abril de 2023).
- [Lap18]** Lapowsky, I. (2018, març 17). Cambridge Analytica Took 50M Facebook Users' Data—And Both Companies Owe Answers. Wired. Recuperat de <https://www.wired.com/story/cambridge-analytica-50m-facebook-users-data>.
- [Mor21]** Moreno, J. (2021, agost 27). Google Estimated To Be Paying \$15 Billion To Remain Default Search Engine On Safari. Forbes. Recuperat de <https://www.forbes.com/sites/johanmoreno/2021/08/27/google-estimated-to-be-paying-15-billion-to-remain-default-search-engine-on-safari>.
- [RCC18]** Rosenberg, M., & Confessore, N., & Cadwalladr, C. (2018, març 17). How Trump Consultants Exploited the Facebook Data of Millions. The New York Times. Recuperat de <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html> (Consultat l'11 d'abril de 2023).
- [Sch18]** Schroepfer, M. (2018, abril 4). An Update on Our Plans to Restrict Data Access on Facebook. Meta Newsroom. Recuperat de <https://about.fb.com/news/2018/04/restricting-data-access> (Consultat l'11 d'abril de 2023).

MATERIAL AUDIOVISUAL.

- [Hin23]** Hindi, R. [Building Web3]. (2023, febrer 23). Homomorphic Encryption and Solving Data Privacy | Founders | Rand Hindi of Zama. Entrevista publicada en YouTube. Recuperat de https://www.youtube.com/watch?v=H6Wr4_eelXI (Consultat l'1 d'octubre de 2023).
- [Kov12]** Kovacs, G. (2012, febrer). Tracking our online trackers. TED Talk. Recuperat de https://www.ted.com/talks/gary_kovacs_tracking_our_online_trackers.
- [KSh20]** KShaikhutdinova. (2020, desembre 15). How Apple and Google Formed One of Tech's Most Powerful Partnerships. Wall Street Journal. Recuperat de <https://www.wsj.com/video/series/wsj-explains/how-apple-and-google-formed-one-of-techs-most-powerful-partnerships/ACED4938-1A00-4031-923F-390E6C6B0B6F>.

COMUNICATS.

- [Zuc18]** Zuckerberg, M. (2018, març 21). I want to share an update on the Cambridge Analytica situation -- including the steps we've already taken and our (...). Facebook. Recuperat de <https://www.facebook.com/zuck/posts/10104712037900071>

LLIBRES.

- [Par11]** Pariser, E. (2011, maig 12). The Filter Bubble: What The Internet Is Hiding From You. Recuperat de <https://dl.acm.org/doi/book/10.5555/2029079>