

Les données que récolte Google - Ch.6

Voici déjà la traduction du sixième chapitre de Google Data Collection, l'étude élaborée par l'équipe du professeur Douglas C. Schmidt, spécialiste des systèmes logiciels, chercheur et enseignant à l'Université Vanderbilt. Si vous les avez manqués, retrouvez les chapitres précédents déjà publiés.

Il s'agit cette fois de comprendre comment Google complète les données collectées avec les données provenant des applications et des comptes connectés des utilisateurs.

Traduction Framalang : Auteurs: Khrys, Piup, Goofy, David_m, Côme, Fabrice, Serici

VI. Données collectées par les applications clés de Google destinées aux particuliers

67. Google a des dizaines de produits et services qui évoluent en permanence (une liste est disponible dans le tableau 4, section IX.B de l'annexe). On accède souvent à ces produits grâce à un compte Google (ou on l'y associe), ce qui permet à Google de relier directement les détails des activités de l'utilisateur de ses produits et services à un profil utilisateur. En plus des données d'usage de ses produits, Google collecte également des identificateurs et des données de localisation liés aux appareils lorsqu'on accède aux services Google. ¹

68. Certaines applications de Google (p.ex. YouTube, Search, Gmail et Maps) occupent une place centrale dans les tâches de base qu'une multitude d'utilisateurs effectuent quotidiennement sur leurs appareils fixes ou mobiles. Le tableau 2 décrit la portée de ces produits clés. Cette section explique comment chacune de ces applications majeures collecte les informations des utilisateurs.

Tableau 2 : Portée mondiale des principales applications Google

Produits	Utilisateurs actifs
Search	Plus d'un milliard d'utilisateurs actifs par mois, 90.6 % de part de marché des moteurs de recherche ²
Youtube	Plus de 1,8 milliard d'utilisateurs inscrits et actifs par mois ³
Maps	Plus d'un milliard d'utilisateurs actifs par mois ⁴
Gmail	1,2 milliard d'utilisateurs enregistrés ⁵

A. Recherche

69. Google Search est le moteur de recherche sur internet le plus populaire au monde ⁶, avec plus de 11 milliards de requêtes par mois aux États-Unis ⁷. En plus de renvoyer un classement de pages web en réponse aux requêtes globales des utilisateurs, Google exploite d'autres outils basés sur la recherche, tels que Google Finance, Flights (vols), News (actualités), Scholar (recherche universitaire), Patents (brevets), Books (livres), Images, Videos et Hotels. Google utilise ses applications de recherche afin de collecter des données liées aux recherches, à l'historique de navigation ainsi qu'aux activités d'achats et de clics sur publicités. Par exemple, Google Finance collecte des informations sur le type d'actions que les utilisateurs peuvent suivre, tandis que Google Flight piste leurs réservations et recherches de voyage.

70. Dès lors que Search est utilisé, Google collecte les données de localisation par différents biais, sur ordinateur ou sur mobile, comme décrit dans les sections précédentes. Google enregistre toute l'activité de recherche d'un utilisateur ou utilisatrice et la relie à son compte Google si cette personne est connectée. L'illustration 13 montre un exemple d'informations collectées par Google sur une recherche utilisateur par mot-clé et la navigation associée.



Illustration 13 : Un exemple de collecte de données de recherche extrait de la page My Activity (Mon Activité) d'un utilisateur

71. Non seulement c'est le moteur de recherche par défaut sur Chrome et les appareils Google, mais Google Search est aussi l'option par défaut sur d'autres navigateurs internet et applications grâce à des arrangements de distribution. Ainsi, Google est récemment devenu le moteur de recherche par défaut sur le navigateur internet Mozilla Firefox ⁸ dans des régions clés (dont les USA et le Canada), une position occupée auparavant par Yahoo. De même, Apple est passé de Microsoft Bing à Google pour les résultats de recherche via Siri sur les appareils iOS et Mac ⁹. Google a des accords similaires en place avec des OEM (fabricants d'équipement informatique ou électronique) ¹⁰, ce qui lui permet d'atteindre les consommateurs mobiles.

B. YouTube

72. YouTube met à disposition des utilisateurs et utilisatrices une plateforme pour la mise en ligne et la visualisation de contenu vidéo. Il attire plus de 180 millions de personnes rien qu'aux États-Unis et a la particularité d'être le deuxième site le plus visité des États-Unis ¹¹, juste derrière Google Search. Au sein des entreprises de streaming multimédia, YouTube possède près de 80 % de parts de marché en termes de visites mensuelles (comme décrit dans l'illustration 14). La quantité de contenu mis en ligne et visualisé sur YouTube est conséquente : 400 heures de vidéo sont mises en ligne chaque minute ¹² et 1 milliard d'heures de vidéo sont visualisées quotidiennement sur la plateforme YouTube. ¹³



Illustration 14 : Comparaison d'audiences mensuelles des principaux sites multimédia aux États-Unis ¹⁴

73. Les utilisateurs peuvent accéder à YouTube sur l'ordinateur (navigateur internet), sur leurs appareils mobiles (application et/ou navigateur internet) et sur Google Home (via un abonnement payant appelé YouTube Red). Google collecte et sauvegarde l'historique de recherche, l'historique de visualisation, les listes de lecture, les abonnements et les commentaires aux vidéos. La date et l'horaire de chaque activité sont ajoutés à ces informations.

74. Si un utilisateur se connecte à son compte Google pour accéder à n'importe quelle application Google via un navigateur internet (par ex. Chrome, Firefox, Safari), Google reconnaît l'identité de l'utilisateur, même si l'accès à la vidéo est réalisé par un site hors Google (ex. : vidéos YouTube lues sur cnn.com). Cette fonctionnalité permet à Google de pister l'utilisation YouTube d'un utilisateur à travers différentes plateformes tierces. L'illustration 15 montre un exemple de données YouTube collectées.



Illustration 15 : Exemple de collecte de données YouTube dans My Activity (Mon Activité)

75. Google propose également un produit YouTube différencié pour les enfants, appelé YouTube Kids, dans l'intention d'offrir une version « familiale » de YouTube avec des fonctionnalités de contrôle parental et de filtres vidéos. Google collecte des informations de YouTube Kids, notamment le type d'appareil, le système d'exploitation, l'identifiant unique de l'appareil, les informations de journalisation et les détails d'utilisation du service. Google utilise ensuite ces informations pour fournir des annonces publicitaires limitées, qui ne sont pas cliquables et dont le format, la durée et le site sont limités.¹⁵

C. Maps

76. Maps est l'application phare de navigation routière de Google. Google Maps peut déterminer les trajets et la vitesse d'un utilisateur et ses lieux de fréquentation régulière (ex. : domicile, travail, restaurants et magasins). Cette information donne à Google une idée des intérêts (ex. : préférences d'alimentation et d'achats), des déplacements et du comportement de l'utilisateur.

77. Maps utilise l'adresse IP, le GPS, le signal cellulaire et les points d'accès au Wi-Fi pour calculer la localisation d'un appareil. Les deux dernières informations sont collectées par le biais de l'appareil où Maps est utilisé, puis envoyées à Google pour évaluer la localisation via son interface de localisation (Location API). Cette interface fournit de nombreux détails sur un utilisateur, dont les

coordonnées géographiques, son état stationnaire ou en mouvement, sa vitesse et la détermination probabiliste de son mode de transport (ex. : en vélo, voiture, train, etc.).

78. Maps sauvegarde un historique des lieux qu'un utilisateur connecté à Maps par son compte Google a visités. L'illustration 16. montre un exemple d'un tel historique ¹⁶. Les points rouges indiquent les coordonnées géographiques recueillies par Maps lorsque l'utilisateur se déplace ; les lignes bleues représentent les projections de Maps sur le trajet réel de l'utilisateur.



Illustration 16 : Exemple d'un historique Google Maps (« *Timeline* ») d'un utilisateur réel

79. La précision des informations de localisation recueillies par les applications de navigation routière permet à Google de non seulement cibler des audiences publicitaires, mais l'aide aussi à fournir des annonces publicitaires aux utilisateurs lorsqu'ils s'approchent d'un magasin ¹⁷. Google Maps utilise de plus ces informations pour générer des données de trafic routier en temps réel.¹⁸

D. Gmail

80. Gmail sauvegarde tous les messages (envoyés et reçus), le nom de l'expéditeur, son adresse email et la date et l'heure des messages envoyés ou reçus. Puisque Gmail représente pour beaucoup un répertoire central pour la messagerie électronique, il peut déterminer leurs intérêts en scannant le contenu de leurs courriels, identifier les adresses de commerçants grâce à leurs courriels publicitaires ou les factures envoyées par message électronique, et connaître l'agenda d'un utilisateur (ex. : réservations à dîner, rendez-vous médicaux...). Étant donné que les utilisateurs utilisent leur identifiant Gmail pour des plateformes tierces (Facebook, LinkedIn...), Google peut analyser tout contenu qui leur parvient sous forme de courriel (ex. : notifications, messages).

81. Depuis son lancement en 2004 jusqu'à la fin de l'année 2017 (au moins), Google peut avoir analysé le contenu des courriels Gmail pour améliorer le ciblage publicitaire et les résultats de recherche ainsi que ses filtres de pourriel. Lors de l'été 2016, Google a franchi une nouvelle étape et a modifié sa politique

de confidentialité pour s'autoriser à fusionner les données de navigation, autrefois anonymes, de sa filiale DoubleClick (qui fournit des publicités personnalisées sur internet) avec les données d'identification personnelles qu'il amasse à travers ses autres produits, dont Gmail ¹⁹. Le résultat : « les annonces publicitaires DoubleClick qui pistent les gens sur Internet peuvent maintenant leur être adaptées sur mesure, en se fondant sur les mots-clés qu'ils ont utilisés dans leur messagerie Gmail. Cela signifie également que Google peut à présent reconstruire le portrait complet d'une utilisatrice ou utilisateur par son nom, en fonction de tout ce qui est écrit dans ses courriels, sur tous les sites visités et sur toutes les recherches menées. » ²⁰

82. Vers la fin de l'année 2017, Google a annoncé qu'il arrêterait la personnalisation des publicités basées sur les messages Gmail ²¹. Cependant, Google a annoncé récemment qu'il continue à analyser les messages Gmail pour certaines raisons ²².