

Histoire de l'origine de Google



[Source : Marguerite Rothe • Le Blogue]

Par Jeff Nesbit

Il y a vingt ans, les services de renseignements américains ont travaillé en étroite collaboration avec la Silicon Valley pour suivre les citoyens dans le cyberspace. Et Google est au cœur de cette histoire d'origine. Certaines des recherches qui ont conduit à la création ambitieuse de Google ont été financées et coordonnées par un groupe de recherche créé par la communauté du renseignement pour trouver des moyens de suivre des individus et des groupes en ligne.

La communauté du renseignement espérait que les plus grands informaticiens du pays pourraient prendre des informations non classifiées et des données sur les utilisateurs, les combiner avec ce qui allait devenir l'Internet, et commencer à créer des entreprises commerciales à but lucratif pour répondre aux besoins de la communauté du renseignement et du public. Ils espéraient diriger la révolution de la super informatique dès le début afin de donner un sens à ce que des millions d'êtres humains faisaient à l'intérieur de ce réseau d'information numérique. Cette collaboration a rendu possible aujourd'hui un état de surveillance de masse public-privé complet.

L'histoire de la création délibérée de l'État moderne de surveillance de masse inclut des éléments de l'origine surprenante, et largement inconnue, de Google. Il s'agit d'une histoire de création quelque peu différente de celle que le public a entendue, et qui explique ce que les cofondateurs de Google, Sergey Brin et Larry Page, ont entrepris de construire, et pourquoi.

Mais il ne s'agit pas seulement de l'histoire de l'origine de Google : C'est l'histoire de l'origine de l'État de surveillance de masse et de l'argent du gouvernement qui l'a financé.

L'histoire : La communauté du renseignement et la Silicon Valley

Au milieu des années 1990, les services de renseignement américains ont commencé à se rendre compte qu'ils avaient une opportunité. La communauté des superordinateurs commençait tout juste à passer du milieu universitaire au secteur privé, grâce à des investissements provenant d'un endroit qui allait devenir la Silicon Valley.

Dès le début, la communauté du renseignement a voulu donner une forme aux efforts de la Silicon Valley afin qu'ils soient utiles à la sécurité intérieure.

Une révolution numérique était en cours : une révolution qui allait transformer le monde de la collecte de données et la façon dont nous donnons un sens à des quantités massives d'informations. La communauté du renseignement voulait donner forme aux efforts de la Silicon Valley en matière de super informatique dès leur création afin qu'ils soient utiles à des fins militaires et de sécurité intérieure. Ce réseau de supercalculateurs, qui deviendrait capable de stocker des téraoctets d'informations, pourrait-il donner un sens intelligent à la trace numérique que les êtres humains laissent derrière eux ?

La réponse à cette question a suscité un grand intérêt dans la communauté du renseignement.

La collecte de renseignements était peut-être leur univers, mais la Central Intelligence Agency (CIA) et la National Security Agency (NSA) avaient pris conscience que leur avenir allait probablement être profondément façonné en dehors du gouvernement. C'était à une époque où les budgets militaires et de renseignement au sein de l'administration Clinton étaient en danger, et où le secteur privé disposait de vastes ressources. Si la communauté du renseignement voulait effectuer une surveillance de masse à des fins de sécurité nationale, il faudrait une coopération entre le gouvernement et les nouvelles sociétés de super informatique.

Pour ce faire, ils ont commencé à prendre contact avec les scientifiques des universités américaines qui étaient à l'origine de cette révolution de la super informatique. Ces scientifiques développaient des moyens de faire ce qu'aucun groupe d'êtres humains assis à des postes de travail de la NSA et de la CIA ne pourrait jamais espérer faire : rassembler d'énormes quantités de données et les interpréter intelligemment.

Une riche histoire du financement de la science par le gouvernement

Il y avait déjà une longue histoire de collaboration entre les meilleurs scientifiques américains et la communauté du renseignement, depuis la création de la bombe atomique et de la technologie des satellites jusqu'aux efforts pour envoyer un homme sur la lune.

L'internet lui-même a été créé grâce à un effort de renseignement.

En fait, l'internet lui-même a été créé grâce à un effort de renseignement : Dans les années 1970, l'agence chargée de développer les technologies émergentes à des fins militaires, de renseignement et de sécurité nationale, la Defense Advanced Research Projects Agency (DARPA), a relié quatre superordinateurs pour traiter des transferts de données massifs. Elle a

confié les opérations à la National Science Foundation (NSF) une dizaine d'années plus tard, qui a fait proliférer le réseau dans des milliers d'universités et, finalement, auprès du public, créant ainsi l'architecture et l'échafaudage du World Wide Web.

La Silicon Valley n'a pas fait exception à la règle. Au milieu des années 1990, la communauté du renseignement a commencé à financer les efforts les plus prometteurs en matière de super calcul dans le monde universitaire, guidant la création d'efforts visant à rendre des quantités massives d'informations utiles à la fois pour le secteur privé et pour la communauté du renseignement.

Ils ont financé ces informaticiens par le biais d'un programme non classifié et très cloisonné, géré pour la CIA et la NSA par de grands contractants militaires et de renseignement. Il s'agissait du projet « Massive Digital Data Systems » (MDDS).

Le projet « Massive Digital Data Systems » (MDDS)

Le MDDS a été présenté à plusieurs dizaines d'informaticiens de premier plan à Stanford, CalTech, MIT, Carnegie Mellon, Harvard, et d'autres dans un livre blanc qui décrivait ce que la CIA, la NSA, la DARPA et d'autres agences espéraient réaliser. La recherche serait en grande partie financée et gérée par des agences scientifiques non classées comme la NSF, ce qui permettrait d'étendre l'architecture au secteur privé si elle parvenait à atteindre les objectifs de la communauté du renseignement.

« Non seulement les activités deviennent plus complexes, mais l'évolution de la demande exige que la communauté du renseignement traite différents types de données ainsi que des volumes plus importants », a déclaré la communauté du renseignement dans son livre blanc de 1993 sur le MDDS. « Par conséquent, la communauté du renseignement joue un rôle proactif en stimulant la recherche dans la gestion efficace de bases de données massives et en veillant à ce que les exigences de la communauté du renseignement puissent être incorporées ou adaptées dans des produits commerciaux. Comme les défis ne sont pas propres à une seule agence, le Community Management Staff (CMS) a chargé un groupe de travail sur les systèmes de données numériques massives (MDDS) de répondre aux besoins et d'identifier et d'évaluer les solutions possibles ».

Au cours des prochaines années, l'objectif déclaré du programme était de fournir plus d'une douzaine de subventions de plusieurs millions de dollars chacune pour faire avancer ce concept de recherche. Les subventions devaient être dirigées en grande partie par la NSF afin que les efforts les plus prometteurs et les plus fructueux puissent être captés sous forme de propriété intellectuelle et constituer la base d'entreprises attirant des investissements de la Silicon Valley. Ce type de système d'innovation public-privé a permis de lancer de puissantes entreprises scientifiques et technologiques comme Qualcomm, Symantec, Netscape et d'autres, et a financé la recherche fondamentale dans des domaines comme le radar Doppler et la

fibre optique, qui sont aujourd'hui au cœur de grandes entreprises comme AccuWeather, Verizon et AT&T. Aujourd'hui, la NSF fournit près de 90 % de l'ensemble du financement fédéral pour la recherche universitaire en informatique.

L'objectif final de la CIA et de la NSA

Les services de recherche de la CIA et de la NSA espéraient que les meilleurs esprits informatiques du monde universitaire pourraient identifier ce qu'ils appelaient les « oiseaux d'une plume » : tout comme les oies volent ensemble en forme de grand V, ou des volées de moineaux font des mouvements soudains ensemble en harmonie, ils ont prédit que des groupes d'humains partageant les mêmes idées se déplaceraient ensemble en ligne. La communauté du renseignement a baptisé sa première réunion d'information non classifiée destinée aux scientifiques « Birds of a Feather »*, et la « Session Birds of a Feather sur l'initiative de la communauté du renseignement en matière de systèmes de données numériques massives » a eu lieu à l'hôtel Fairmont de San José au printemps 1995.

La communauté du renseignement a baptisé son premier briefing non classifié destiné aux scientifiques le briefing « birds of a feather » (les oiseaux d'une plume).

Leur objectif de recherche était de suivre les empreintes digitales à l'intérieur du réseau d'information mondial en pleine expansion, qui était alors connu sous le nom de World Wide Web. Un monde entier d'informations numériques pourrait-il être organisé de manière à ce que les demandes faites par les humains à l'intérieur d'un tel réseau soient suivies et triées ? Leurs demandes pourraient-elles être reliées et classées par ordre d'importance ? Pourrait-on identifier les « oiseaux de plume » dans cette mer d'informations afin que les communautés et les groupes puissent être suivis de manière organisée ?

En travaillant avec de nouvelles sociétés de données commerciales, leur intention était de suivre des groupes de personnes partageant les mêmes idées sur Internet et de les identifier à partir des empreintes digitales qu'ils avaient laissées, un peu comme les scientifiques utilisent les empreintes digitales pour identifier les criminels. Tout comme « les oiseaux de plume s'assemblent », ils ont prédit que les terroristes potentiels communiqueraient entre eux dans ce nouveau monde global et connecté – et qu'ils pourraient les trouver en identifiant des modèles dans cette énorme quantité de nouvelles informations. Une fois ces groupes identifiés, ils pouvaient alors suivre leurs traces numériques partout.

Sergey Brin et Larry Page, les jeunes prodiges de l'informatique

En 1995, l'une des premières et des plus prometteuses bourses MDDS a été attribuée à une équipe de recherche en informatique de l'université de Stanford, qui travaillait depuis dix ans avec des bourses de la NSF et de la

DARPA. L'objectif principal de cette subvention était « l'optimisation des requêtes très complexes qui sont décrites à l'aide de l'approche « query flocks » ». Une deuxième subvention, celle de la DARPA-NSF, la plus étroitement associée à l'origine de Google, faisait partie d'un effort coordonné pour construire une bibliothèque numérique massive utilisant l'Internet comme épine dorsale. Les deux subventions ont financé les recherches de deux étudiants diplômés qui ont fait des progrès rapides dans le classement des pages web, ainsi que dans le suivi (et la compréhension) des requêtes des utilisateurs : les futurs cofondateurs de Google, Sergey Brin et Larry Page.

Les recherches menées par Brin et Page dans le cadre de ces subventions sont devenues le cœur de Google : des personnes utilisant des fonctions de recherche pour trouver précisément ce qu'elles voulaient dans un très vaste ensemble de données. La communauté du renseignement, cependant, a vu un avantage légèrement différent dans leurs recherches : Le réseau pouvait-il être organisé de manière si efficace que les utilisateurs individuels pouvaient être identifiés et suivis de manière unique ?

Ce processus est parfaitement adapté aux objectifs de la lutte contre le terrorisme et de la sécurité intérieure : Les êtres humains et les groupes de même sensibilité qui pourraient représenter une menace pour la sécurité nationale peuvent être identifiés de manière unique en ligne avant de nuire. Cela explique pourquoi la communauté du renseignement a trouvé les efforts de recherche de Brin et Page si attrayants ; avant cette époque, la CIA utilisait largement les efforts de renseignement humain sur le terrain pour identifier les personnes et les groupes susceptibles de constituer une menace. La possibilité de les suivre virtuellement (en conjonction avec les efforts sur le terrain) allait tout changer.

C'était le début de ce qui, dans quelques années, allait devenir Google. Les deux responsables de la communauté du renseignement chargés de diriger le programme ont rencontré régulièrement Brin au fur et à mesure de l'avancement de ses recherches. Il a également été l'auteur de plusieurs autres documents de recherche résultant de cette subvention MDDS avant que lui et Page ne partent pour former Google.

Les subventions ont permis à Brin et à Page de faire leur travail et ont contribué à leurs percées en matière de classement des pages web et de suivi des requêtes des utilisateurs. Brin ne travaillait pas pour la communauté du renseignement – ni pour personne d'autre. Google n'avait pas encore été constitué en société. Il n'était qu'un chercheur de Stanford profitant de la subvention accordée par la NSA et la CIA dans le cadre du programme MDDS non classifié.

Laissé de côté dans l'histoire de Google

L'effort de recherche du MDDS n'a jamais fait partie de l'histoire des origines de Google, même si le chercheur principal de la subvention MDDS a spécifiquement nommé Google comme résultant directement de leur recherche :

« Sa technologie de base, qui lui permet de trouver des pages bien plus précisément que les autres moteurs de recherche, a été partiellement soutenue par cette subvention », a-t-il écrit. Dans un document de recherche publié qui inclut certains des travaux essentiels de Brin, les auteurs font également référence à la subvention de la NSF qui a été créée par le programme MDDS.

Au lieu de cela, chaque histoire de création de Google ne mentionne qu'une seule subvention fédérale : la subvention NSF/DARPA pour les « bibliothèques numériques », qui a été conçue pour permettre aux chercheurs de Stanford de faire des recherches sur l'ensemble de la Toile mondiale stockée sur les serveurs de l'université à l'époque. « Le développement des algorithmes de Google a été effectué sur divers ordinateurs, principalement grâce au projet de bibliothèque numérique de Stanford financé par la NSF/DARPA-NASA », explique par exemple l'Infolab de Stanford à propos de son origine. De même, la NSF ne fait référence qu'à la subvention pour les bibliothèques numériques, et non à la subvention MDDS également, dans sa propre histoire de l'origine de Google. Dans le célèbre document de recherche, « The Anatomy of a Large-Scale Hypertextual Web Search Engine », qui décrit la création de Google, Brin and Page remercie la NSF et la DARPA pour la subvention accordée à Stanford pour les bibliothèques numériques. Mais la subvention du programme MDDS de la communauté du renseignement – spécifiquement conçue pour la percée sur laquelle Google a été construit – s'est effacée dans l'obscurité.

Google a déclaré par le passé qu'il n'était pas financé ou créé par la CIA. Par exemple, lorsque des histoires ont circulé en 2006 selon lesquelles Google avait reçu des fonds de la communauté du renseignement pendant des années pour aider à la lutte contre le terrorisme, la société a déclaré au fondateur du magazine Wired, John Battelle, que « les déclarations relatives à Google sont complètement fausses ».

La CIA a-t-elle directement financé le travail de Brin et Page, et donc créé Google ? Non. Mais Brin and Page faisait-il des recherches précisément pour répondre aux attentes de la NSA, de la CIA et de la communauté du renseignement, avec l'aide de leurs subventions ? Absolument.

La CIA et la NSA ont financé un programme non classifié et cloisonné conçu dès le départ pour favoriser la création d'un site ressemblant presque à Google.

Pour comprendre cette importance, il faut considérer ce que la communauté du renseignement essayait de réaliser en accordant des subventions aux meilleurs esprits informatiques du monde universitaire : La CIA et la NSA ont financé un programme non classifié et cloisonné conçu dès le départ pour stimuler le développement de quelque chose qui ressemble presque exactement à Google. La recherche révolutionnaire de Brin sur le classement des pages en suivant les requêtes des utilisateurs et en les reliant aux nombreuses recherches effectuées – identifiant essentiellement les « oiseaux de plume » – était en grande partie l'objectif du programme MDDS de la communauté du renseignement. Et Google a réussi au-delà de leurs rêves les plus fous.

L'héritage durable de la communauté du renseignement dans la Silicon Valley

Ces dernières années, les préoccupations relatives à la protection de la vie privée en milieu numérique, à l'intersection entre la communauté du renseignement et les géants de la technologie commerciale, se sont accrues. Mais la plupart des gens ne comprennent toujours pas à quel point la communauté du renseignement s'appuie sur les plus grandes entreprises scientifiques et technologiques du monde pour son travail de lutte contre le terrorisme et de sécurité nationale.

Les groupes de défense de la liberté civile font part de leurs préoccupations en matière de protection de la vie privée depuis des années, notamment en ce qui concerne le Patriot Act. « Adopté à la hâte 45 jours après le 11 septembre au nom de la sécurité nationale, le Patriot Act a été le premier des nombreux changements apportés aux lois sur la surveillance qui ont permis au gouvernement d'espionner plus facilement les Américains ordinaires en élargissant le pouvoir de surveiller les communications téléphoniques et les courriels, de collecter les dossiers bancaires et de crédit et de suivre les activités d'Américains innocents sur Internet », explique l'ACLU. « Alors que la plupart des Américains pensent qu'il a été créé pour attraper les terroristes, le Patriot Act transforme en fait les citoyens ordinaires en suspects ».

Lorsqu'on leur pose la question, les plus grandes entreprises de technologie et de communication – de Verizon et AT&T à Google, Facebook et Microsoft – affirment qu'elles ne proposent jamais délibérément et de manière proactive leurs vastes bases de données sur leurs clients aux agences fédérales de sécurité et d'application de la loi : Elles affirment qu'elles ne répondent qu'aux citations à comparaître ou aux demandes déposées en bonne et due forme en vertu du Patriot Act.

Mais même un bref coup d'œil dans les archives publiques récentes montre qu'il existe un tapis roulant de demandes constantes qui pourraient saper l'intention derrière cette promesse de protection de la vie privée. Selon les dossiers de demande de données que les entreprises mettent à la disposition du public, au cours de la dernière période de référence, entre 2016 et 2017, les autorités locales, étatiques et fédérales à la recherche d'informations liées à la sécurité nationale, à la lutte contre le terrorisme ou à des préoccupations criminelles ont émis plus de 260 000 citations à comparaître, ordonnances de tribunal, mandats et autres demandes légales à Verizon, plus de 250 000 demandes de ce type à AT&T, et près de 24 000 citations à comparaître, mandats de recherche ou ordonnances de tribunal à Google. Les demandes directes liées à la sécurité nationale ou à la lutte contre le terrorisme ne représentent qu'une petite fraction de ce groupe de demandes, mais la procédure juridique du Patriot Act est désormais tellement routinière que les entreprises ont chacune un groupe d'employés qui s'occupent simplement du flux de demandes.

Ainsi, la collaboration entre la communauté du renseignement et les grandes entreprises commerciales scientifiques et technologiques a connu un succès

foudroyant. Lorsque les agences de sécurité nationale doivent identifier et suivre des personnes et des groupes, elles savent à qui s'adresser – et le font fréquemment. C'était l'objectif au départ. Il a réussi peut-être plus que quiconque aurait pu l'imaginer à l'époque.

(Traduction DeepL)

Pour avoir accès aux liens hypertexte dans l'article (anglais), suivre le lien ci-dessous :

Google's true origin partly lies in CIA and NSA research grants for mass surveillance