Chapitre conclusif : “Power, Knowledge, and Responsibility: The Ethics of Ontological Systems in the Age of AI”
Ontologies, Données & Pouvoir Algorithmique

L’avènement du Big Data et la multiplication des données disponible a entraîné l’apparition de nouveaux défis : comment gérer cette nouvelle masse de données sans précédent ? Comment l’exploiter ? Comment connecter ensemble des données hétérogènes issues de sources et de formats différents ? Gérer leur intégration continue ? Ces défis ont été en partie résolus par des innovations techniques permettant de mieux gérer les flux de données, mieux les exploiter.
De même cette profusion de données a apporté de nombreux potentiels : analyse prédictive des données pour prédire des ventes, repérer les points de données irréguliers (pouvant par exemple correspondre à des fraudes) et les traiter de façon adéquate.
Néanmoins force est de constater que le problème du traitement de la donnée hétérogène n’est pas résolu. En pratique, la plupart des données restent encore isolées en silo, au sein même de la même entreprise. Par exemple la gestion des clients, la gestion des fiches de paies sont prises en charge par des services logiciels différents et qui ne communiquent pas entre eux. Or un comptable pourrait souhaiter avoir accès aux données via le même système sans pour autant devoir extraire les informations de suites logicielles différentes. C’est un constat assez frappant sur le monde du logiciel de façon plus générale : les innovations techniques issues du Big Data ont permis de mieux gérer les flux de données et de les canaliser de façon pertinente pour une seule application précise, mais sans grande souplesse. La solution logicielle prend en charge sa mission (ex : gérer les fiches de paie) et le fait de façon efficace avec les bonnes données mises à jour, mais les données et leurs traitement ne sont pas exploités plus largement. De même si les modalités de calcul de paie évoluent, le système n’est pas assez flexible pour prendre en compte les évolutions de législation et nécessite une mise à jour adaptée. Les outils logiciels développés en général ne permettent pas cet agilité qu’à l’esprit humain de pouvoir connecter deux informations isolées en un constat pertinente (from knoledge to intelligence) en raison de ces contraintes techniques. L’essentiel du travail des informaticiens est passé à orienter un certain type de données vers l’application monotâche ce qui limite de facto les potentialités de la solution développée. Mais tant qu’elle fonctionne, pas de problème. L’inconvénient est que avec cette manière de procéder, on atteint assez rapidement les limites de l’automatisable : peut être automatisé toute procédure pouvant être explicitement décrite par une suite d’étapes et de transformations à partir de certaines données prévues en entrée. Cela implique que l’on sache exactement quel type d’information on veut découvrir (approche top down). Or dans beaucoup de cas d’usage, on ignore les phénomènes qui nous intéressent jusqu’à ce qu’une approche quantitative ou qualitative appropriée nous l’ait fait faite apercevoir. Les flux de données évoluent de manière dynamique et sont rarement standardisés. Il faut de plus éviter la compartimentation de plusieurs sources de données qui s’éclairent potentiellement l’une l’autre. Ce qui est la clé de tout n’est pas forcément l’échantillon le plus significatif. Même les systèmes d’IA les plus avancés atteignent leurs limites lorsqu’il s’agit de naviguer dans une masse de données hétérogènes. C’est pourquoi le traitement en masse de la donnée hétérogène est un verrou technique important à faire céder pour pouvoir repousser les limites du domaine de l’automatisable et exploiter pleinement la quantité de données incessamment produite par les personnes ou les institutions. Résoudre ce défi d’ingénierie pour l’instant le plus souvent effleuré limite actuellement

le potentiel des solutions intégrant l'intelligence artificielle, qu'elle soit prédictive ou générative. Le résoudre ouvre la voie à de nombreux cas d'usages et réalisations.

Toutefois, derrière ce discours progressiste visant à « enfin donner accès à la machine à la richesse des données représentant le savoir et l'expérience humains » se cachent des enjeux éthiques qui sont souvent éclipsés par d'autres débats, sur la délégation de l'IA à la prise de décision par exemple. Cette volonté de dompter la donnée hétérogène peut tout d'abord être rapproché de la volonté cartésienne de voir l'homme comme maître et possesseur de la nature. La nature, réduite ici à un ensemble de points de données la modélisant deviendrait un objet que l'informatique pourrait analyser, prédire, ce qui optimiserait l'emprise technique qu'à l'homme sur elle.
D'autre part, cette volonté d'exploiter les données reste une volonté de contrôle, tout d'abord de contrôle de processus matériels pour améliorer les rendements de production (analyse de chaînes de production industrielles), mais l'histoire de la révolution industrielle a montré (Foucault) que l'optimisation des moyens de production a très souvent été associée à une augmentation du contrôle sur les individus eux mêmes. Or le même outil technique qui permet d'agréger les informations sur la production d'un téléphone aux quatre points du globe peut avec une facilité déconcertante être utilisé pour agréger des informations diverses sur des personnes. Il s'agit derrière des mêmes mécanismes techniques de gestion de la donnée hétérogène. Se pose alors la question de la vie privée des individus si leurs données issues de différentes sources comme par exemple des institutions gouvernementales, dess données collectées par des plateformes de ventes, donnée libres d'accès sur le web (OSINT) sont analysées par des algorithmes qui recensent, classent et étiquètent. Et l'ingénierie des connaissances est fortement impliquée dans ce processus de collecte. En organisant des données éparses grâce à des ontologies, cette branche de l'informatique pourrait faciliter les dérives autoritaires. Transformer la donnée brute en connaissance exploitable est déjà un acte de pouvoir, car il s'agit de choisir ce qui compte, ce qui est lié, ce qui doit être vu ensemble. Aussi la manière de concevoir les ontologies traduit une certaine vision du monde qui est celle des auteurs ou des commanditaires de l'algorithme. Ce qui pourrait être vu comme le simple prolongement technique des algorithmes de recommandation largement utilisés actuellement pourrait, orienté par des pouvoirs étatiques ou autres, devenir un outil de surveillance et de contrôle de masse. Il est donc nécessaire de réfléchir à la façon dont les défis techniques rencontrés actuellement en informatique s'intersectent avec des problèmes éthiques qu'ils pourraient soulever.

On reviendra d'abord sur la notion de donnée personnelle pour voir que son statut légal et juridique en fait un objet bien plus complexe et protéiforme que la définition restrictive que l'on en pourrait avoir. C'est précisèrent la capacité de relier ensemble des informations qui transforment une donnée en donnée personnelle. On étudiera ensuite en prenant l'exemple de l'entreprise américaine Palantir comment une exploitation de la donnée hétérogène (problème essentiellement technique) résolue à l'aide d'ontologies peut amener à des difficultés éthiques avant de tenter de définir un cadre de réflexion autour de ces pratiques.

I) La place fluctuante des données personnelles dans la collecte de masse

La collecte et l'analyse de données en masse offre des potentiels économiques non négligeables, notamment sur l'analyse des marchés, l'optimisation de réseaux de transport ou de processus industriels… Mais même dans ces perspectives purement commerciales émergent déjà des problématiques éthiques. Quelles données peut-on légitimement collecter pour mener ses analyses ? A priori, l'usage de toute information disponible sur le web ouvert peut être considérée comme fair game. Mais que penser des données, souvent plus personnelles partagées sur les réseaux sociaux ? Si des protections empêchent à des personnes non inscrites et, si le profil est privé, non connectées à la personne d'intérêt empêchent normalement à des sources tierces d'accéder aux contenus publiés par les personnes, les réseaux sociaux (Twitter notamment) commercialisent un accès direct à ces contenus. Que penser de même des banques de données client le plus souvent collectées à force de cookies, à l'insu des utilisateurs ? A l'heure du Big Data, la définition et la protection des données personnelles est un enjeu d'importance croissante. Il convient de comprendre l'évolution du statut de la donnée personnelle pour mieux comprendre les enjeux juridiques et éthiques que cette notion pose actuellement.

1) le nouveau paradigme de la donnée personnelle

En 2025, un constat est que chaque être humain crée constamment de la donnée. Quand on utilise un produit numérique, se connecte à une plateforme, cette utilisation laisse une trace sur des réseaux qui même si elle semble anodine peut être exploitée (bracelets fitbit des gardes du corps de personnalités politiques utilisés pour déterminer leur domicile). Chaque déplacement que l'on fait via un véhicule, chaque achat, chaque formulaire rempli, message posté constitue autant de supports qui contiennent des informations potentiellement exploitables sur leur auteur. Chaque personne est un flux de données continuel qui vient se glisser dans l'océan. Avant l'avènement du big data, cela ne posait pas de problème ; seul un détective privé aurait pu collecter les fragments et les assembler pour effectuer du profilage. Mais avec la numérisation des achats, des services publics et des moyens de communication, toutes ces productions d'une part se retrouvent sous le même format, un format numérique, et d'autre part les outils qui permettent de les collecter et de les relier sont de plus en plus performants. C'est en ce sens que, parmi les nombreuses citations tarte à la crème sorties sur le Big Data, Humby déclaire que « data is the new oil ». L'idée que la donnée est un produit transformable. On peut en faire un carburant (entraîner des modèles), un composant (base de données) si on dispose des outils permettant son extraction et sa transformation.

Cette mise à disposition de contenus[1] créés ou portant sur une personne soient directement exploitables pose des questions éthiques. La première portant évidemment sur la propriété de ces données. La question se règle le plus souvent dans les petites lignes des contrats d'utilisation que les utilisateurs négligent de lire atttentivement, ou qui, du moins outre Atlantique peuvent être modifiées à la discrétion de l'entreprise. Cela conduit à des ambiguïtés[2]

1 Notion de « vorhanden » dans la réflexion générale sur la technique.

2 « Twitter's Terms of Service Do Not Permit Third Parties to Commercially Use Content » : mais si on ne vend pas directement la donnée ? Si on s'en sert juste pour classer un profil ?

La donnée personnelle puisqu'elle peut être transformée en profit, devient elle-même quelque chose de commercialisable. L'example le plus marquant concerne peut être les tests de séquençage d'ADN autorisés aux US : des entreprises en faillite sont autorisées à revendre ces données pour payer leurs dettes. Quoi de plus personnel que la cartographie de l'ADN qui est ce qui fait de chacun une personne unique ?

Arthur Nielsen, pionnier américain de l'analyse de marché, a jeté les bases de cette industrie en transformant les comportements des consommateurs en informations exploitables pour les entreprises. Aujourd'hui, cette logique est poussée à l'extrême par les "**data brokers**[3]" ou courtiers en données, qui collectent, analysent et revendent des informations personnelles à grande échelle.
Ces entreprises, souvent méconnues du grand public, telles qu'Acxiom, Experian ou Equifax, compilent des milliers de données sur chaque individu : habitudes de consommation, déplacements, interactions en ligne, et bien plus encore. Leur objectif principal est de prédire et d'influencer le comportement des consommateurs, en vendant ces profils détaillés à des annonceurs, des institutions financières ou des compagnies d'assurance. L'impact social de ces pratiques est significatif. Par exemple, les données agrégées par les data brokers sont utilisées pour évaluer la solvabilité des individus, influençant ainsi l'accès au crédit aux assurances ou même à l'emploi. (surtout aux US où avoir un bon credit score est nécessaire pour contracter un prêt immobilier) Cette situation rappelle le rôle des agences de notation comme Moody's, qui évaluent la fiabilité financière des États, avec des conséquences directes sur leur économie.

En France, bien que la culture de la protection des données soit plus ancrée, notamment grâce au RGPD, les activités des data brokers restent peu visibles et peu comprises du grand public. Pourtant, leur influence sur la vie quotidienne est réelle et soulève des questions éthiques majeures concernant la vie privée et la transparence des algorithmes décisionnels.
Radio Franc Face à ces enjeux, des initiatives émergent pour encadrer ces pratiques. Aux États-Unis, le Consumer Financial Protection Bureau (CFPB) a proposé de nouvelles réglementations visant à limiter la vente d'informations personnelles sensibles par les data brokers, les assimilant à des agences de crédit soumises à des obligations strictes.

Une des conclusions de ce nouveau paradigme de l'information est que si c'est gratuit, vous êtes vraisemblablement le produit. Que faire des lors pour protéger ses données ? Quel est le cadre législatif mis en place pour définir et protéger les données personnelles ?

2) Différents cadres juridiques pour protéger la donnée personnelle

Le cadre législatif peine souvent à s'adapter à l'évolution de la technique. Soit la réglementation fige l'innovation à un instant donné et devient inadaptée à suivre son avancée, soit

---

3 https://www.lemonde.fr/pixels/article/2025/02/12/donnees-personnelles-en-vente-libre-les-data-brokers-une-industrie-hors-de-controle_6543025_4408996.html
https://medium.com/@siavash.alamouti/exposing-the-world-of-data-brokers-a-serious-threat-to-global-prosperity-d825dcef6295
https://www.sup.org/books/sociology/data-cartels

elle tarde trop pour encadrer les pratiques les plus néfastes liées à une technologie. Ce problème est d'autant plus accentué lorsque l'objet du différend n'est lui même pas évident à définir.

En France, la CNIL définit par exemple une donnée personnelle une donnée qui permet d'identifier une personne directement ou indirectement.[4]

« Une personne physique peut être identifiée :
•directement (exemple : nom et prénom) ;

•indirectement (exemple : par un numéro de téléphone ou de plaque d'immatriculation, un identifiant tel que le numéro de sécurité sociale, une adresse postale ou courriel, mais aussi la voix ou l'image).

L'identification d'une personne physique peut être réalisée :

•à partir d'une seule donnée (exemple : nom) ;

•à partir du croisement d'un ensemble de données (exemple : une femme vivant à telle adresse, née tel jour et membre dans telle association).

Cette définition est elle-même très vague. Parfois un nom et prénom ne permettent pas une identification directe. D'autre part le croisement de données anodines (« employé d'une boîte vivant dans telle ville ayant tel âge… ») permet d'identifier clairement un unique individu.
Il convient donc de préciser plus clairement ce que l'on entend par identification.
Cela implique qu'un travail de profilage préalable a été réalisé, ou on évalue les capacités d'évaluation à l'aune d'une figure hypothétique qui aurait le profil de tous les individus vivant sur terre.

Le cadre de la RGPD[5] vient préciser la première définition partielle de 2018. ce nouveau texte a pour mérite de reconnaître le statut problématique des individus comme étant des producteurs de données et souligne également l'existence de données à caractère sensible (« Ce sont les données qui révèlent l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses, philosophiques ou l'appartenance syndicale, les données génétiques, les données biométriques, les données concernant la santé, les données qui concernent la vie sexuelle et l'orientation sexuelle ».)
La RGPD définit un statut particulier à ces données qui ne peuvent être collectées sans un consentement explicite. Ces données nécessitent aussi d'être traitées et stockées de façon particulières (interdiction de partage à des entités tierces, protection pour limiter les risques en cas de piratage. Ce statut de donnée sensible suit la donnée durant tout le cours de son existence, de sa collecte à sa destruction.
« *Traitement* » est le terme générique employé dans le RGPD pour désigner une **opération quelconque** sur des données personnelles. En effet, le RGPD s'applique aux traitements de données personnelles.
« Selon la définition très large qu'en donne le RGPD, il s'agit de toute opération effectuée sur une donnée personnelle, telle que : La collecte, l'enregistrement, la structuration, le stockage, l'extraction, la modification / rectification, la consultation l'utilisation, la publication, la communication la diffusion ou toute autre forme de mise à disposition, l'interconnexion, l'effacement et la destruction »

4 https://www.cnil.fr/fr/definition/donnee-personnelle
5 https://datalegaldrive.com/rgpd-tout-savoir/rgpd-donnees-personnelles/

Cette définition intéressante de la donnée personnelle comme donnée identifiable touche au problème nodal du big data et des KGs : la donnée n'est jamais isolée ; c'est précisément l'inférence (profiling, liaison) permise par des outils de Big Data qui la rend sensible.

Cette définition européenne de la donnée sensible et personnelle n'est pas unanimement partagée ce qui conduit régulièrement à des contentieux, tel que le procès intenté à la CNIL à la start-up new yorkaise Clearview AI[6]. Cette entreprise avait construit, à partir de milliards de photographies collectées sur le web, un outil de reconnaissance faciale.
Clearview AI a enfreint le protocole concernant les données sensibles sur plusieurs aspects :
* Le fait de pouvoir associer un visage à un profil tombe à la fois sous le coup de la donnée personnelle et de la donnée sensible (biométriques).
* la collecte a eu lieue sans obtenir le consentement, ni même informer les personnes concernées.

Face à la demande de la CNIL de supprimer les données des résidents français, l'entreprise rétorque que il est impossible à partir d'une photo de savoir si une personne est française ou non et que l'image est accessible sur les moteurs de recherche. De plus, officiant aux US ils n'auraient pas à se soumettre à la RGPD. Cette entreprise avait été financée par Peter Thiel. Le compte rendu de la CNIL statue que « Le fait qu'une donnée soit **publiquement accessible ne lui fait par perdre son caractère personnel** et les principes sur les données personnelles restent applicables. » et que pour que son règlement s'applique, il suffit que le traitement de la donnée soit lié au suivi du comportement de personnes en Europe. L'entreprise refuse à ce jour encore de se conformer aux lois européennes. Si l'exemple semble anecdotique, il traduit un difficile alignement commun sur de bonne pratiques de protection des données et préfigure de plus vastes mécanismes de surveillance.

3) Les moyens techniques de préserver l'anonymisation

Une des contraintes posées par le nouveau cadre législatif européen autour de la donnée est que si une donnée est sensible et ou personnelle, elle doit être stockée de façon anonymisée. C'est à dire que si un acteur malveillant parvenait à accéder au périphérique de stockage, il serait dans l'incapacité de relier l'information à une personne physique. Cette donnée ne serait donc plus par définition une donnée personnelle. De même, pour exploiter des données dans des domaines sensible (médecine, juridique) pour entraîner des modèles, il faut que ces données aient été rendues anononymes, pour éviter par exemple que le modèle ne « recrache » des informations personnelles qui étaient présentes dans ses données d'entraînement[7].

Mais comment faire techniquement pour anonymiser des données ? Cela revient à faire précisément l'inverse des solutions de Big data qui aggrègent des données hétérogènes éparpillées. On appelle anonymisation l'ensemble des techniquest « de manière à rendre impossible, en

6 https://www.lemonde.fr/pixels/article/2022/10/20/reconnaissance-faciale-la-cnil-condamne-clearview-ai-a-une-amende-de-20-millions-d-euros_6146699_4408996.html
7 https://arxiv.org/abs/2302.00539

pratique, toute identification de la personne par quelque moyen que ce soit et de manière irréversible »[8].

Il existe diverses techniques de complexité et d'efficacité variable.

* la pseudonymisation consiste à empêcher l'attribution directe d'une donnée à une personne sans information supplémentaire. Cela consiste le plus souvent en remplaçant des données directement identifiable (couple nom prénom, date de naissance, numéro de sécurité sociale) par des alias ou des id. Toutefois, dans les faits, la réidentification de la personne reste souvent possible en recoupant des données avec d'autres données extérieures (par exemple, la présence de l'âge des individus peut permettre de ré-identifier très facilement les personnes centenaires). Ces opérations ne sont pas irréversibles contrairement à l'anonymisation dans sa définition la plus pure.

* la génralisation consiste, quand on a pas besoin de l'information exacte à remplacer une information précise (date de naissance, adresse...) par une information plus générale qui ne permet pas la ré-identification (tranche d'âge, département de résidence…)

* la randomisation consiste à échange, ou permuter de façon aléatoire les attributs correspondant à plusieurs profils. Il faut toutefois conserver une clé ou un mécanisme permettant de relire les données au risque de les rendre inexploitables. Je précise ; les données restent exploitables pour des analyses quantitatives si la randomisation est faite de sorte à ne pas dégrader l'équilibre des groupes. En revanche les données sur les profils individuels deviennent faussées.

Ces techniques ont toutes un point faible majeur qui est que même si les données stockées dans le système ne permettent plus d'identifier une personne, peut être que recouper la base de données avec une autre base de données (publique, d'une autre entreprise…) permettrait de le faire. On parle d'attaques par reconstruction ou de ré-identification[9]. Et il y a aucun moyen de s'assurer que cela ne soit pas le cas puisqu'on ne connaît pas les bases de données extérieures.

Faces aux limites des techniques traditionnelles d'anonymisation, d'autres méthodes algorithmiques ont été développées. L'une d'entre elle est la confidentialité différentielle (differential privacy, technique inventée par Cynthia Dwork entre autres[10]) qui vise à pouvoir permettre de profiter des informations déductibles d'une grande masse de données sensibles (statistiques médicales par exemple) tout en s'assurant que l'identification de données sur une personne précise (un patient) soit absolument impossible. Il faut donc tenir un équilibre délicat entre s'assurer que chaque donnée soit anonymisée mais que le dataset dans son ensemble de même composition statistique que le dataset non anonymisé

. L'idée est d'introduire dans le dataset (aggrégé) un bruit contrôlé de manière à créé de l'anonymat pour les individus. Par exemple sur un dataset de données médicales, si on s'intéresse à l'âge des personnes et veut calculer la moyenne, l'écart type, alors on peut par des opérations retirer quelques années à des lignes du tableau, en ajouter à d'autres… de telle sorte que les indicateurs statistiques ne soient pas modifiés. Evidemment c'est bien plus compliqué que cela car il faut tenir en compte de l'interdépendance entre les attributs. Par exemple le lien entre l'âge et la pathologie,

8 https://www.cnil.fr/fr/technologies/lanonymisation-de-donnees-personnelles
9 https://arstechnica.com/tech-policy/2009/09/your-secrets-live-online-in-databases-of-ruin/ voir l'anecdote sur un malheureux gouverneur américain
10 https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf

le genre, le lieu de naissance …). Le risque avec cette approche est de dégrader l'équilibre des données et de les rendre inutilisables pour le cas d'usage qu'on voulait traiter.

Pour éviter cela, il faut « garder le bruit sous contrôle ». Dans le calcul des modifications faites au dataset, il existe un facteur epsilon qui mesure l'équilibre entre utilisabilité et fiabilité des données et anonymat obtenu. Plus epsilon est petit, plus on ajoute de bruit et plus la confidentialité est assurée, et inversement. Plus le dataset est grand et plus il est facile d'atténuer le tradeoff. La méthode permet aussi de quantifier le changement maximal des valeurs d'intêrét du dataset.

La définition formelle de la differential privacy est la suivante.

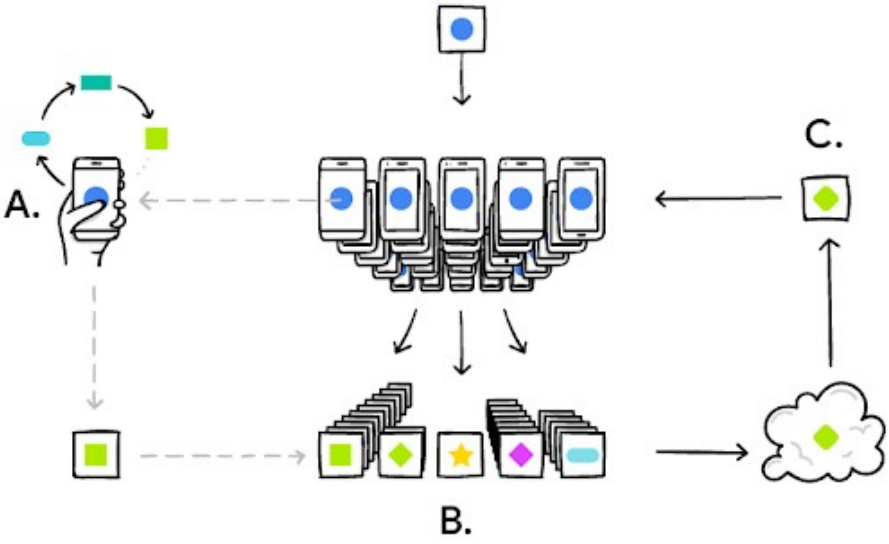
Soit D et D' deux datasets n'ayant comme différence que une seule nouvelle entrée dans D', et que l'on interagit avec les données selon un mécanisme M, alors on dit que M est epsilon différentiellement confidentiel si pour toute sortie de M, la probabilité que cette sortie soit observée ne diffère pas de plus de exp(epsilon) entre les deux scénari.

$$\forall D \text{ and } D' \text{ that differ in one person's data } \forall x: \mathbb{P}[M(D) = x] \leq exp(\varepsilon) \cdot \mathbb{P}[M(D') = x]$$ [11]

M peut être tout et n'importe quoi : une requête dans une base de données, une transaction dans un système de fichier … mais aussi un modèle d'IA. Si les paramètres d'un modèle sont observés avec des probabilités qui ne diffèrent pas de exp(epsilon) l'une de l'autre, alors la procédure d'entraînement conserve l'anonymat. Le fait qu'une observation individuelle soit présente dans les données d'entraînement ne modifie pas le comportement du modèle. La difficulté est qu'il faut arriver à prouver ça … pour tous les points de donnée individuels du dataset.

Cette approche a été appliquée en IA pour s'assurer de l'anonymat dans les données d'entraînement et que les paramètres ne stockent pas des données qu'ils ne sont pas censés stocker. Il existe par exemple une version modifiée de l'algorithme de descente de gradient stochastique (DP-SGD *Differentially-Private Stochastic Gradient Descent) q*ui modifie l'optimisation du réseau de neurone de sorte que l'accès aux graients des paramètres se fasse de façon à préserver les principes de confientialité différentielle. Si l'entraînement respecte à chaque étape cette confidentialité, alors le modèle obtenu au final la respecte aussi.

Une autre approche en IA qui protège l'intégrité des données est de procéder à du « federated learning » quand on entraîne un modèle. Au lieu de devoir stocker toutes les données d'entraînement sur la même machine, ce qui expose à des risques de capture des données ou de connexion de points de données ensemble, on peut laisser les données d'entraînement sur les machines où elles étaient initialement, ce qui empêche de relier les données éclatées sur plusieurs machines. Chaque appareil doit



11 https://medium.com/pytorch/dif

télécharger la version en cours du modèle, voir quels changements le passage des batches causerait, et transmettre ces changements au serveur principal pour qu'ils soient appliqués au modèle. Pour rendre cela possible, il a fallu modifier les algorithmes d'entraînement, notamment SGD. Google se sert par exemple de cette approche pour améliorer les suggestion de complétion sur les claviers. Selon les options que les utilisateurs sélectionnent les premiers choix proposés sont modifiés grâce aux données communes des utilisateurs.

La définition de données privées et sensibles a grandement évolué avec l'arrivée du Big Data, de même que les moyens algorithmiques permettant de la préserver. Toutefois, cette confidentialité des données est mise à mal au fur et à mesure que les capacités et la latence des logiciels permettant d'agréger les données augmente : plus ils seront performants et plus les données pseudonymisées seront identifiées. Or c'est précisément toutes les techniques d'ingénierie des connaissances qui sont en partie responsables de l'augmentation de la capacité des logiciels à agréger les données ensemble. On étudiera un cas pratique, qui est la façon dont le logiciel Palantir utilise la notion d'ontologie pour procéder à l'agrégation de données hétérogènes.

II) Les ontologies, nerf de la guerre du renseignement : le cas de Palantir

L'ingénierie des connaissances peut être utilisée pour traiter des données sensibles et personnelles, notamment dans les domaines du renseignement et du marketing, ce qui peut engendrer des préjudices encore mal circonscris ou pris en charge par la loi. On étudiera pour illustrer ce point l'entreprise américaine Palantir dont le but est d'aider ses clients à exploiter une masse informationnelle diffuse pour aider à la prise de décision. Née avec un objectif de défense nationale, l'entreprise a au fil des ans développé une gamme de services et de logiciels adaptés à la défense et aux univers régaliens, mais aussi décliné pour les entreprises civiles. La particularité de Palnatir est de collaborer étroitement avec leurs clients, les aidant à développer une « ontologie » de leur domaine (j'utilise des guillemets à dessein, car il restera a déterminer si la définition d'une ontologie telles qu'utilisées dans les logiciels de Palantir respecte ou non la définition traditionnelle qu'on a élaboré en NLP), ontologie qui est ensuite exploitée pour analysée des flux de données et les exploiter. On présentera l'histoire de cette entreprise et les produits qui sont commercialisés avant d'étudier comment les ontologies et les techniques d'ingénierie des connaissances, de façon plus large sont au coeur de leur fonctionnement.

1) Palantir, objectif, philosophie et histoire

L'entreprise Palantir a été fondée en 2003 par Peter Thiel, Joe Lonsdale, Stephen Cohen et Alex Karp, dont le premier et le dernier restent très actifs dans le milieu tech et la politique américains. Ils méritent sans doute une courte biographie :

* Peter Thiel est un investisseur germano américain ayant étudié la philosophie et le droit à Stanford où il rencontre René Girard. Il se convertit rapidement en investisseur et profite du boom du web. Faisant porter ses efforts sur la cryptographie, il a l'idée de développer une plateforme permettant de façon sécurisée de gérer un portefeuille numérique en 1999 : PayPal ! Après la revente de PayPal à ebay, Thiel s'est concentré sur son nouveau projet, Palantir. Il arbore clairement des vues conservatrices, libertariennes et nationalistes. Il est l'auteur de quelques essais tel que « The diversity Myth » (critique du politiquement correct et des politiques de diversité). Il affirme dans un essaie de 2009 : "[I]] no longer believe that freedom and democracy are compatible".

* Stephen Cohen est un ingénieur américain formé à Stanford qui aurait selon la légende conçu le prototype de Palantir en deux mois[12]. Il s'est formé en IA et en NLP auprès de Andrew Ng, notamment.

* Alex Karp : nommé comme une des 100 personnes les plus influentes du monde en 2025 par Forbes. A fait des études de philosophie (élève d'Habermas!) et est surtout connu pour son rôle de CEO de Palantir. the embodiment of a new kind of Silicon Valley billionaire: an unashamed techno-nationalist who evangelizes Western power." (Time magazine) Il a récemment (Février 2025) publié un livre 'The Technological Republic: Hard Power, Soft Belief, and the Future of the West ». L'argument principal en est que les entreprises de technologie, basées dans la silicon valley notamment, se sont détournées de leurs objectif initial qui était l'avancée du savoir humain, appliqué à l'amélioration de nos conditions de vies, pour ne se concentrer plus que sur la recherche aveugle du profit. Il appelle à une réévaluation de ces objectifs et de plus amples partenariats gouvernementaux.[13]

* Joe Lonsdale : Ingénieur et investisseur également passé par PayPal, il a travaillé sur le prototype de Palantir avant de s'écarter du groupe pour plus d'investissements… Le reste n'est pas glorieux.

La plupart des fondateurs sont issus de la « paypal mafia », nom donné à un groupe d'investisseurs ayant fait fortune au moment de la mise en place de modes de paiement numériques. L'entreprise prend son nom dans un artefact fictif des romans de Tolkien, les Palantiri ou « pierres de visions », pierres permettant la communication entre elles et qui avaient été données aux hommes par les elfes pour relier chaque citadelle. Or, dans la saga du Seigneur des Anneaux, ces pierres ont été utilisées par Sauron pour corrompre ceux qui les ont utilisés. L'outil est donc ambigü et dangereux comme le reconnaît Lonsdale « there is a warning built into the name »[14]. Le projet derrière Palantir est né comme conséquence des attentats du 11 septembre que les services secrets

12 https://web.archive.org/web/20150503204351/https://www.washingtonian.com/articles/people/killer-app/indexp3.php

13 https://www.ft.com/content/8ea36422-2f65-4a14-93be-b7b4d38362e3
https://www.lesechos.fr/tech-medias/intelligence-artificielle/video-alex-karp-le-milliardaire-derriere-palantir-2164182

14 In the fantasy realm of Tolkien, the palantir seeing stones were made thousands of years ago by the elves of Valinor in the Uttermost West, and gifted to their friends around the world. These stones could communicate to help see the past and understand the future, and were used to secure the world by overcoming forces of evil. Unfortunately, in future ages, they eventually fall into the wrong hands and are used for corrupt purposes — their images distorted to pervert goals, and their power harnessed to further evil. https://www.quora.com/Did-Palantirs-founders-consider-the-ethical-implications-of-their-work-Do-they-have-thoughts-about-the-consequences-of-what-they-built/answer/Joe-Lonsdale?ref=thediff.co

américains ont été incapables d'anticiper. Le but est de proposer une solution logicielle qui vienne aider les enquêteurs et les analystes dans leur mission. Le constat était que les services américains collectent énormément de données mais qu'il n'est pas facile de les exploiter, de rassembler ensemble celles qui doivent l'être pour pouvoir avancer dans des enquêtes. Ce travail est en général réalisé par des analystes professionnels qui étudient les données « à la main », mais ont une capacité d'absorption et de connexion limitée. Ainsi l'enjeu du big data se retrouvait directement lié à un enjeu de sécurité nationale. Principe (Mais c'était en 2003!) que l'IA ne saurait seule à repérer des mesures adaptables et que l'expertise humaine était nécessaire. Le but du logiciel devient donc de présenter les données à l'utilisateur pour lui permettre de faire des connections, accélérer son pouvoir de décision.
« intelligence augmentation »

Les fondateurs de Palantir se proposaient comme mission de « reduce terrorism while preserving civil liberties ». En plus de cette philosophie civique, la création de Palantir est liée à une certaine philosophie entreprenariale valorisant la prise de risque et la mise à distance de problèmes éthiques délicats[15].

L'entreprise a connu une histoire mouvementée. Depuis sa création en 2003/4, un corps réduit d'employés s'acharnait à développer un prototype et a obtenir des contrats gouvernementaux, ce qui s'est produit au début des années 2010.
Palantir Metropolis, Quantity Analysis Tool a été utilisé avec succès pour repérer les fraudes dans les agences gouvernementales américaines puis pour supprimer les silos d'information entre les agences américaines et centraliser leurs données. À partir de là, l'entreprise a connu une forte expansion et a diversifié son activité dans le secteur privé. Ayant quitté la Silicon Valley pour s'installer à Denver, Colorado, l'entreprise fait partie des valeurs reconnues de la tech américaine avec un revenu voisinant les 3 Milliards, 4000 employés.

Que penser de cette trajectoire ? Elle est assez inhabituelle dans le monde de la technologie.[16]
* un objectif politique affirmé mais ambigu : « défendre les libertés civiles » et pas de produit intéressant le grand public.[17]
* un succès économique avec une indexation au S&P 500, des bénéfices record.
* de nombreux scandales et une opinion publique très mitigée

De 2016 à 2020, mauvaise réputation à cause des scandales sur la protection des données.
* une culture qui met en avant « l'importance de l'intenstité », hérité de l'esprit de la « mafia Paypal » : travailler dur, ne pas accepter la défaite, une bonne dose de compétitivité
* un arrière plan philosophique fort derrière les enjeux techniques. Peter Thiel and Alex Karp were philosophy grads « The overall 'vibe' of the company was more of a messianic cult than a normal software company » « There's something to this correlation: by making the company about something other than making money (civil liberties; AI god) you attract true believers from the start »
* une stratégie de terrain avec des équipes découpées en deux : des ingénieurs qui vont à la rencontre des clients dans des industries critiques pour développer des solutions qui résolvent leurs

15https://www.8vc.com/resources/lessons-from-peter-thiel

16https://medium.com/@nabeelqu/reflections-on-palantir-52433cf95439

17 there was a technological solution to the challenge of balancing public safety and civil liberties — a "Hegelian" aspiration

problèmes, et des ingénieurs travaillant sur des produits qui améliorent les premières esquisses (scalabilité, performances) et créent des outils qui automatisent les tâches répétitives. Les produits (Foudry notamment) ce sont petit à petit construits en amalgamant des outils effectuant une partie du travail de traitement des données. (service company → product company pivot)
Toutes ces caractéristiques font de Palantir un ovni dans le monde de la tech. Le principal étant que la technologie développée n'est absolument pas grand public mais adressée à de grandes institutions. Il convient de mieux comprendre l'offre de Palantir.

2) la déclinaison des produits

Palantir a construit son offre logicielle pendant plus de 20 ans. L'étude des produits proposés, mis en rapport avec l'histoire de l'entreprise permet de voir comment une technologie de traitement en masse de la donnée hétérogène a été construite.

a) Les produits phares

Quand on consulte le site web de Palantir, on constate que 4 produits principaux sont présenté : Palantir Gotham, Foundry, Apollo et AIP. Il convient de constater que ce découpage de l'offre est en réalité assez récent et on reviendra plus en détail sur la construction de ces produits.

Les quatre produits mis en avant par Palantir différent selon leur socle technique et le public visé :

* Gotham : « intelligence » et contre terrorisme. Compte comme client le département de la défense des US entre autres. Son développement a pris la suite en 2008 du produit Métropolis. Il renforce notamment les capacités des analystes par l'intégration de multiples sources de données, y compris des données géospatiales, les connecte, et effectue des traitements par IA. Petit à petit, cet outil s'est spécialisé à des fins militaires permettant l'acquisition de cibles en temps réels, le suivi de l'évolution des mouvements de troupes. Une partie des fonctionnalités d'analyse pure de Métropolis, originalement destinée à des clients étatiques a été déplacée dans des versions de Palantir Foundry[18] pendant que Gotham sert à des missions purement étatiques, comme par exemple l'optimisation de la « kill chain »lors d'opérations militaires.

* Foundry : plateforme complète pour l'intégration de données des entreprises (« The Ontology-Powered Operating System for the Modern Enterprise »). Palantir Foundry, comme Gotham d'ailleurs, n'est « que » l'organisation dans une plateforme unique de tous les modules nécessaires au traitement de la donnée. Le produit a pour tôle de connecter ensemble toutes les données d'une entreprise en une plateforme unique (data lake). La plupart du temps, les entreprises disposent d'une masse écrasante de données qui sont réparties en silo (chaque département / équipe a des données très spécifiques à sa zone de compétence) mais qui ne sont pas reliées entre elle ce qui limite leur valeur (la valeur d'une donnée se calcule en fonction de l'usage pratique qu'on peut en faire). La plateforme se veut comme une boîte à outil qui permet de réaliser tous les câblages, de la collecte des données depuis les différents silos jusque son emploi dans des applications.

18 https://www.youtube.com/watch?v=rxKghrZU5w8

La force de Palantir est d'avoir su adapter cette unique plateforme à des besoins métier très différents, comme la lutte contre la fraude, l'optimisation de Supply Chain, la gestion de données médicales… Je vous laisse lire la liste complète sur leur site. L'idée est que le produit s'est adapté et a développé des modules spécifiques pour traiter certains cas d'usage (par exemple le traitement de données satellites).

* Apollo : facilite l'intégration continue de données (CI /CD) dans tous les environnements.
C'est le système de CD qui permet de déployer les deux précédents logiciels précédents et d'autres encore. L'idée est de pouvoir déployer un logiciel sur n'importe quel environnement machine (la différence de système d'exploitation peut par exemple déjà être un souci énorme quand on installe un logiciel). Et encore on parle de machines et systèmes d'exploitation main stream et pas de machines sécurisées et spécialisées comme certains ordinateurs militaires… Fort de leur expérience à installer leur logiciel dans des environnements très différents (et l'installation peut parfois occuper une partie non négligeable du temps de travail sur un projet), les équipes ont développé un programme pour pouvoir déployer facilement leurs logiciels quel que soit l'environnement de réception, pour pouvoir effectuer les mises à jour de façon plus facile.

* AIP : artificial Intelligence Platform[19]

Ce dernier produit de Palantir, plutôt une extension des précédents prévoir l'intégration des LLM dans des réseaux opérés de façon privée de sorte à assurer la confidentialité des données. L'utilisateur peut configurer des actions employant des LLMs.
Par exemple, en cas d'alertes (nouvelles informations apportées par les connecteurs), l'IA peut analyser la donnée, classer son urgence et proposer des actions à suivre. L'humain garde le contrôle en validant ou invalidant la marche à suivre proposée par le logiciel et a la possibilité de modifier les prompts pour personnaliser la solution à son besoin.
L'utilisateur a la possibilité d'ajouter à l'IA des outils (fonction de code…) permettant au LLM de chercher des informations ou modifier des données à la demande. Ces actions par IA peuvent être intégrées dans les pipeline de traitement à différentes étapes
Finalement on peut dire que AIP laisse créer à l'utilisateur des agents via une interface GUI et ensuite permet d'intégrer ces agents dans des applications.

=> Le point commun de ces produits est qu'ils se présentent comme des systèmes d'exploitation (Operating systems) et proposent un univers d'exploration complet : l'utilisateur charge ses données, les traite, les exploite, écris des rapports et envoie des commandes depuis le logiciel même.

b) Les modules indépenants

Au fur et à mesure de son travail avec les clients, les équipes de Palantir ont construit des briques logicielles ayant une fonction spécifique qui ont ensuite été combinées intelligemment pour donner des produits : Gotham et Foundry.

S'intéresser au rôle et au fonctionnement de ces produits permet de mieux cerner ce que fait vraiment l'entreprise pour ses clients.

19 https://www.palantir.com/platforms/aip/

Comme le présente de façon très intéressante un ancien employé[20], chaque développement d'une fonctionnalité était conçue pour résoudre un besoin particulier :
« Need to bring in data from SAP or AWS? Here's Magritte (a data ingestion tool). Need to visualize data? Here's Contour (a point and click visualization tool). Need to spin up a quick web app? Here's Workshop (a Retool-like UI for making webapps) »
Le nom de ces applications a petit a petit été effacé au cours de leur intégration dans Foundry, mais comme on y reviendra sous peu, ces modules constituent en fait des blocs nécessaires au traitement de la donnée.

c) la stratégie de verticalisation

Palantir a suivi une stratégie originale de développement de produit par rapport à de nombreuses boîtes de tech, les plaçant à mi chemin entre un éditeur de logiciel et une forme de consulting. Souvent une boîte de tech construit un produit pour viser un marché, ou répond à un appel d'offre pour créer un produit répondant au besoin de l'utilisateur. Ces besoins sont clairement définis dans un cahier des charges et le vainqueur de l'appel d'offre se doit de suivre un calendrier (souvent avec des retards) pour livrer un produit intégrant de plus en plus de fonctionnalités et testé au fur et à mesure par les utilisateurs finaux. Mais ne disposant pas de produit, pendant longtemps la stratégie de Palantir était de se présenter comme solveurs des problèmes d'une institution en réorganisant son traitement de la donnée. Au début de l'entreprise, presque tous les employés étaient ingénieurs et se répartissaient en deux groupes : des experts qui étaient envoyés en équipe directement auprès des clients. Ils discutaient de vive voix avec eux, étudiaient la mission à effectuer, les données à collecter et leur support, les verrous techniques pour réussir la mission, et parvenaient rapidement à construire un prototype adapté au problème, qui était ensuite amélioré en collaboration avec les équipes de l'institution d'accueil. L'autre partie des ingénieurs ont pour fonction de reprendre ces prototypes créées par les experts présents sur le terrain, et de les améliorer en les rendant :
* plus scalables et robustes
* plus propre, mieux organisé, plus facile à installer
* si possible, plus génériques, capable de traiter le même problème mais dans d'autres entreprises.
Cela consiste à éviter le risque d'« overfitting » de la solution et de petit à petit construire un outil complet et versatile pour le traitement de données.[21]
Un très bon exemple de cette stratégie est le développement d'un produit pour Airbus pour optimiser la fabrication des A 350[22]. Des experts de Palantir ont été délégués directement dans les usines de Toulouse pour observer la construction des avions et la chaîne d'approvisionnement et sont restés des mois auprès des équipes d'Airbus.

C'est avec cette stratégie que Palantir a gagné peu à peu d'autres domaines d'activité que le domaine régalien : l'entreprise décroche un contrat dans un secteur d'activité spécifique et utilise à la fois la gamme de module existants et en développe de nouveaux adaptés pour traiter les problèmes du client mais aussi monter en compétences sur les besoins de ce secteur d'activité. Puis dans un second temps, une solution adaptée à ce secteur d'activité, mais pas liée aux spécificités

20 https://medium.com/@nabeelqu/reflections-on-palantir-52433cf95439
21 « This two-pronged model made for a powerful engine. Customer teams were often small (4–5 people) and operated fast and autonomously; there were many of them, all learning fast, and the core product team's job was to take those learnings and build the main platform. »
22 https://www.palantir.com/impact/airbus/

d'un client unique est mise au point. Par exemple, en collaboration avec Airbus, Palantir a mis au point Skywise une plateforme permettant aux compagnies aériennes de centraliser les informations sur leurs vols et qui est désormais adoptée par de nombreuses compagnies concurrentes.
Cette façon de procéder déclenche ce que l'on appelle un network effect ("a phenomenon whereby a product or service gains additional value as more people use it."[23]) Plus le nombre d'utilisateurs augmente, plus le service est conseillé et se répandde client en client, de secteur en secteur.

3) Les ontologies au coeur du produit ?

Le logiciel Palantir repose sur une notion abstraite d'ontologie (« ontology ») qui ne recoupe pas exactement la définition traditionnelle d'une ontologie en informatique. L'ontologie dans Palantir est un moyen de représenter un objet qui joue un rôle dans les processus métier. L'ontologie dans Palantir ne repose pas sur des termes mais sur des objets, les rapprochant plutôt d'une classe en Programmation Orientée Objet. En effet, chaque objet dispose d'un schéma[24]. Le logiciel permet de créer les classes d'objets, d'ajouter leurs propriétés, de créer des sous classes et des instances de ces objets. On peut considérer qu'on objet d'une ontologie est l'équivalent d'un objet en python, ou encore d'une observation, d'une ligne d'un dataset[25]. Les utilisateurs déclarent ensuite des relations (canoniques) pouvant connecter certains types d'objets, et des évènements. Un service permet de pouvoir choisir des noms de champs consistants d'un objet à un autre (dates de naissance …) entre plusieurs objets pour pouvoir effectuer des recherches par agrégation.

C'est à partir de ces « ontologies » que les utilisateurs peuvent construire leurs applications. Une fois les objets et les relations créées, les utilisateurs peuvent mettre en place des règles de raisonnement (flag de certains objets satisfaisant certaines propriétés, création d'un nouvel objet…), des fonctions (calcul de totaux à partir d'objets de la même famille satisfaisant les mêmes propriétés) et créer des applications mettant en valeur certains types d'objets, leurs relations, et ce de façon dynamique.

L'étape suivante une fois que l'ontologie est créée (la plupart du temps manuellement par des spécialistes du métier) et intégrée dans des applications, la tâche la plus complexe reste de la peupler à partir des données réelles qui peuvent être structurées (mais dans un format ne correspondant pas forcément à la structure de l'objet) ou non structuré. C'est là où l'emploi du mot ontologie prend tout son sens et rejoint celui que l'on en faisait dans le cadre de l'ingénierie des connaissances. Il faut donc créer des chemins d'intégration depuis les sources de données vers les objets pertinents et si besoin effectuer les transformations nécessaires, suivant le chemin classique d'une Data Preparation. Pour ce faire, Palantir a mis au point un outil « Pipeline Builder » [26] La collecte des données se fait à partir du chargement de fichier, ou de connecteurs, et les utilisateurs peuvent concevoir des chaînes de traitement adaptées aux données collecter et décrire (construire à

23 https://www.reddit.com/r/PLTR/comments/x35mhv/palantir_is_not_a_product_its_an_enterprise/?rdt=64534

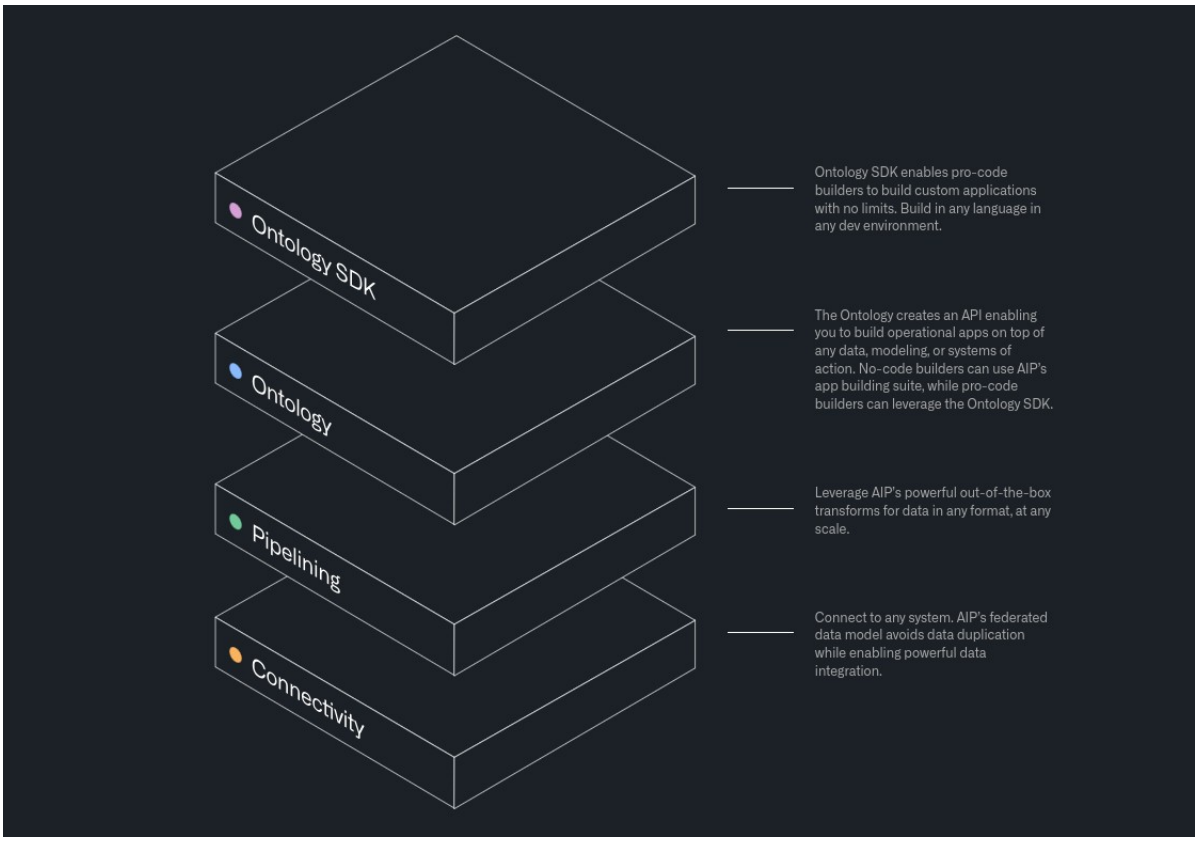
24 https://www.palantir.com/docs/foundry/object-link-types/object-types-overview

25 « The definition of an object type in the Ontology is analogous to that of a dataset, while the definition of an object is analogous to that of a row in the dataset »
https://www.palantir.com/docs/foundry/object-link-types/object-types-overview

26 https://www.palantir.com/docs/foundry/pipeline-builder/overview

l'aide de blocs de logiciel en low code / avec du code / co consrtruire avec un LLM) les transformations nécessaires pour aboutir de la donné brute à l'objet.

L'otologie est donc la fondation de l'offre de la boîte. La pyramide des services offerts commence avec l'ontologie[27] :



Elle peut être créée avec le module spécialisé (ontology manager) ou via des scripts (SDK de création d'ontologie). Une fois l'ontologie finalisée, une API est générée automatiquement pour pouvoir peupler l'ontologie et requêter les objets (construction d'applications interagissant avec l'ontologie) et les utilisateurs peuvent mettre au point des pipelines (avec aide de l'IA)[28].

Si ce n'est pas dit clairement et que la conception palantirienne de l'ontologie diffère sérieusement de celle développée en linguistique, Palantir a finalement construit toute son activité autour de l'ingénierie des connaissances et propose une solution visant à épargner à ses clients tout le travail difficile pour pouvoir exploiter ses données… au risque parfois de rendre les institutions clientes complètement dépendantes du logiciel lui-même et de faire de Palantir un creuset collectant une quantité problématique de données sensibles…

III) Vers un totalitarisme numérique ?

Utilisées à bon escient, les techniques d'ingénierie des connaissance ont le potentiel d'exploiter des données jamais mises en contact jusqu'alors et de faire gagner du temps et de l'argent à de nombreuses organisations (d'optimiser des procédés de fabrication, faire de la

27 https://www.palantir.com/aip/developers/#modular-interoperable
28 https://www.youtube.com/watch?v=vSJ5H00V7Es&t=114s

maintenance prédictive). Ces solutions, conformément aux ambitions initiales des fondateurs peut aussi permettre de repérer les risques, de fraude financière notamment. Et des limites morales commencent à se poser : quelles données est-il légitime d'exploiter pour permettre de repérer des potentielles actions terroristes à venir ? À partir de quand une donnée est-elle trop privée ou sensible pour pouvoir légitimement être exploitée ? Optimiser les processus est très bien quand on parle de fabrication d'avions, de la production des pièces détachées à la livraison, mais quand on parle de l'identification d'une cible à son exécution ? On peut dès lors se questionner sur les zones d'ombre des solutions d'ingénierie des connaissance.

1) les scandales emblématiques

Malgré ses succès techniques et financiers, l'histoire de Palantir a été ponctuée d'un nombre non négligeable de scandales qui remettent en question la légitimité morale des ambitions du groupe et un usage sûr des outils développés. On citera trois affaire intéressantes et symptomatiques de l'ambiguïté fondamentale de cette démarche.

* Le projet Thémis ou la chasse aux sorcières

Des cadres de Palantir ont été impliqués dans le projet Thémis[29] qui consistait à faire du profilage des opposants à la chambre du commerce.
Le projet a fini par être exposé à la presse, arrêté, et la direction de Palantir a déclaré ne pas être au courant. Cette équipe s'en était notamment prise aux membres de wikileaks[30].

* Utiliser la datascience pour sortir de l'UE : Les Cambridge Analytica (2016)

Une société fondée par un certain Christopher Wyllie, Cambridge Analytica, a utilisé des techniques de datascience dans le but d'influencer des scrutins politiques importants, notamment l'élection de Trump en 2016 ou le Brexit[31]. Des techniques de datascience sont exploitées (analyse de sentiment, aggrégation et clusterings) pour identifier sur les réseaux sociaux des profils d'indécis politiques, la tranche de la population qui fait le plus souvent basculer une élection. Des techniques de NLP permettent à moindre coût d'identifier les thèmes importants pour la personne pour pouvoir lui envoyer du contenu sur ce thème, « personnalisé » et les poussant à voter le parti soutenu. En plus du caractère controversé d'une telle entreprise, pour pouvoir mettre en pratique ce plan, Cambridge Analytica a collecté des données de plus de 50 millions d'utilisateurs de Facebook sans leur consentement pour effectuer du profilage[32]. Palantir a fini par reconnaître après plusieurs démentis, qu'un de ses employes avait eu accès à ces données problématiques.[33]

---

29 https://www.washingtonian.com/2012/01/31/killer-app/ Palantir and HBGary Federal teamed up with a third intelligence contractor, Berico Technologies, to provide information on groups and individuals deemed hostile to the US Chamber of Commerce. The law firm Hunton & Williams first approached Palantir about the work, which was to include reconnaissance of various Web sites and social media in order to build dossiers on the chamber's opponents. Operating under the name Team Themis, the companies would set up an analysis cell to provide the law firm with intelligence about "adversarial entities and networks of interest," according to a proposal the team drew up. Palantir would "serve as the foundation for all of the data collection, integration, analysis, and production efforts.

30 https://www.jacobsilverman.com/p/preservation-over-cause-remembering

31 https://www.amazon.fr/ing%C3%A9nieurs-du-chaos-Giuliano-Empoli/dp/2709664062 Voire sur la question entre autres l'excellent livre de Giuliano Da Empoli

32 https://www.marianne.net/monde/le-tres-inquietant-pouvoir-de-palantir-la-boite-melee-au-scandale-facebook-la-nsa-et-la-dgsi

33 https://www.nytimes.com/2018/03/27/us/cambridge-analytica-palantir.html

Le logiciel aurait permis le microciblage politique via des ontologies psychométriques et Cambridge Analytica aurait pu aider Palantir à monter en compétence sur le profilage.

* ImmigrationOs : un logiciel pour expulser. Collaboration polémique entre Palantir & ICE

Un des débats les plus médiatisés autour de Palantir est son association avec l'agence ICE (Immiration and Customs Enforcment, agence américaine chargée de la gestion de l'immigration et des douane). Palantir avait signé pendant le premier mandat de Trumpt un contrat pour développer un logiciel permettant à cette agence de centraliser ses données. Cela avait entraîné une vive réaction publique et des manifestations devant le siège de l'entreprise[34]. Dans le cadre de ce premier contrat, deux outils ont été livrés, ICM (Investigative Case Management) qui permet aux agents de frontière de tracer les liens de filiation entre les nouveaux arrivants et les personnes présentes sur le territoire américain. Le second outil, FALCON sert à mener des raids sur les lieux de travail et arrêter des personnes migrantes qui n'ont pas de documents. Malgré le caractère humainement problématique de l'usage des outils et leur implication dans les expulsions aux États-Unis (2017–2020) Palantir a accepté le contrat et fourni les outils de fusion et traçage d'identités Un nouveau contrat de 30 millions de dollar vient d'être signé avec la nouvelle administration Trump[35].

L'ingénierie des connaissances peut donc avoir des domaines d'application fortement problématiques. Mais il faut aussi souligner que le caractère problématique n'est pas seulement issu des usages de l'outil, mais aussi de la démarche intellectuelle sous jacente.

2. Le fantasme de l'ontologie parfaite

*L'*ingénierie des connaissances vise à construire des représentations structurées et exploitables du savoir humain, et cela a le plus souvent été fait sous la forme d'ontologie, que ce soit une ontologie au sens classique (avec des termes, des catégories) ou au sens d'objet comme dans Palantir.

L'ambition de l'ingénierie des connaissance serait donc de capturer la complexité du réel dans ce formalisme, derrière les ontologies il y a une prétention à représenter *toute* la réalité. Or en IC deux démarches s'affrontent, qui peinent toutes deux à vraiment capturer les objets :

Une démarche top down où le logiciel / l'expert définit son objet. C'est celle de Palantir où il faut en amont construire un objet et ses propriétés. Cela implique un acte de définition qui limite l'objet à la liste de propriétés que l'on rentre dans le logiciel et est forcément partiel. On configure ensuite de façon forcée les flux de données pour qu'ils viennent nourrir cette ontologie.

---

34 https://truthout.org/articles/immigrant-rights-activists-renew-push-against-palantir-to-cancel-ice-contract/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+TRUTHOUT+%28t+r+u+t+h+o+u+t+%7C+News+Politics%29&amp
https://action.mijente.net/petitions/disrupt-the-tech-talent-pipeline-tell-palantir-to-drop-its-contracts-with-ice-anteaters

35 https://multinationales.org/fr/a-chaud/actualites/comment-palantir-le-geant-de-la-big-data-collabore-a-la-chasse-aux-migrants-aux

Une démarche bottom up consiste à l'inverse à partir de la donnée non structurée de découvrir les objets pertinents ce qui implique beaucoup moins de stabilité dans les objets découverts. De plus c'est une tâche de NLP qui a toujours été plus qu'épineuse (Cf Chapitre 2 & 3) et que l'IA générative ne sait pas traiter de façon conséquente.

La volonté d'avoir des objets informatiques représentant le monde de façon fidèle est donc un leurre. Cette question d'avoir une représentation du monde assez fine pour en saisir toute la complexité est présentée comme le chaînon manquant pour faire progresser l'IA au-delà du plateau des données que l'on a quasiment atteint[36].

Il existe en effet un décalage entre la représentation d'un objet dans l'ontologie et l'objet du monde réel ce qui fait écho à la notion de « semantic gap » en informatique. En génie logiciel, le **semantic gap** désigne la différence entre la façon dont un utilisateur humain conçoit un concept, et la façon dont il est modélisé ou représenté en code. Ce problème très ancien du logiciel n'est pas totalement réglé par l'ingénierie des connaissances.

Le risque d'avoir un modèle du monde qui soit trop stéréotypé est d'induire un certain nombre de fausses corrélations qui sont construites à partir d'un accès partiel à l'information. Cela conduit par exemple à *naturaliser* des constructions sociales (e.g., criminalité, radicalité). Palantir a été utilisé par le service de police de la Nouvelle Orléans pour organiser des raids préventifs et réduire la criminalité. Mais le logiciel identifiait certaines zones de la ville comme étant à risque stigmatisant la population y vivant et conduisant à une boucle : plus de raids dans une zone entraîne plus d'arrestations, ce qui encourage le logiciel à considérer la zone comme effectivement dangereuse… Le logiciel qui n'est pas tellement plus qu'un énorme tableau excell ne peut pas déconstruire des corrélations qu'il trouve dans la vue partielle des données.

Toutes ces observations pourraient nous permettre de conclure à l'existence d'une violence ontologique qui découle d'un emploi trop poussé et aveugle des techniques d'ingénierie des connaissances. Gayatri Chakravorty Spivak a introduit le concept de « violence épistémique » dans son essai influent *Can the Subaltern Speak?* (1988). Elle y décrit la manière dont les structures coloniales de savoir ont réduit au silence les voix des subalternes en imposant des cadres de pensée occidentaux, rendant ainsi leurs expériences et connaissances inintelligibles ou illégitimes. Cette violence est « épistémique » car elle opère au niveau de la production et de la légitimation du savoir. Utiliser des systèmes qui catégorisent les objets crée un cadre de pensée statique dans lequel tout ne fonctionne que si ce cadre est respecté scrupuleusement, et sans écart. Ceux qui maîtrisent ce cadre (les modalités de sa construction) maîtrisent le système, alors que les autres individus subissent le cadre qui est imposé à leur existence.

L'usage d'ontologies conduit aussi à la volonté de tout classifier pour forcer les objetts du monde à rentrer dans les cadres prévus. Geoffrey C. Bowker et Susan Leigh Star, dans leur ouvrage *Sorting Things Out: Classification and Its Consequences* (1999), explorent comment les systèmes de classification influencent la société. Ils montrent que ces systèmes, souvent perçus comme neutres, peuvent en réalité marginaliser certains groupes en imposant des catégories qui ne reflètent pas leurs réalités. Cette marginalisation peut être considérée comme une forme de violence ontologique, car elle affecte la manière dont les individus sont perçus et se perçoivent eux-mêmes.

36 https://innovations.fr/lia-world-model-la-cle-vers-une-ia-de-niveau-humain/

Comme mentionné explicitement par les créateurs de Palantir au moment de justifier le nom de l'entreprise, un outil dépend des usages qu'on en fait et peut bien utiliser, apporter des bénéfices là où il peut être très néfastes entre de mauvaises main, mais le sujet devient vraiment sérieux lorsqu'une entreprise place la notion de sécurité au centre de ses objectifs. La frontière entre sécurité et surveillance devient alors très fine, surtout quand des acteurs politiques sont impliqués. Les collusions entre technologie et pouvoir se multiplie : les institutions de pouvoir finissent par emprunter leur vocabulaire et leurs processus aux outils qu'elles ont utilisé et dont elles sont devenues dépendants, tandis que les entreprises de tech gagnent une influence croissante sur les institutions. Cette ambition est clairement affichée quand Alex Karp déclare par exemple souhaiter que Palantir devienne « le système d'exploitation du gouvernement des Etats Unis »[37] et collabore en ce moment avec le Department of Government Efficiency (DOGE) pour créer une « super API » permettant de centraliser les données financières de tous les citoyens américains.

Ce parcours nous aura conduit des promesses techniques de l'ingénierie des connaissances à ses implications politiques et éthiques les plus profondes. Dans un monde saturé de données, l'ambition de structurer l'information pour en extraire du sens n'est pas neutre : elle engage une vision du monde, des hiérarchies de valeur, et des logiques de pouvoir.

Les ontologies, loin d'être de simples artefacts techniques, sont devenues des instruments centraux de gouvernementalité algorithmique. Leur usage dans des systèmes comme ceux de Palantir révèle une tension fondamentale : celle entre **l'optimisation des processus** et **le risque de surveillance**, entre **intelligence augmentée** et **violence ontologique**.

Face à ces enjeux, une posture critique est nécessaire : non pour rejeter en bloc les apports de ces technologies, mais pour interroger leurs conditions de conception, d'usage et de légitimation. Il s'agit de reconnaître que tout système de classification produit des inclusions… mais aussi des exclusions ; et que la prétention à modéliser le réel risque toujours de le simplifier, de l'appauvrir, ou de l'assujettir.

Dès lors, penser des ontologies responsables, c'est poser la question de **qui conçoit**, **pour qui**, **avec quels objectifs**, et **selon quelles valeurs**. La technique ne suffit plus : il faut y réinjecter du débat, du droit, de la politique — et, fondamentalement, de l'humain.

37« become the US government operatin system » youtube.com/watch?v=DZ95Gmvg_D4https://www.youtube.com/watch?v=rxKghrZU5w8