https://farid.ps/articles/industrialization_of_profiling/fr.html

Des Personnes d'Intérêt aux Populations : L'Industrialisation du Profilage

L'émergence du profilage — du scrutin manuel ciblé de « personnes d'intérêt » spécifiques à la surveillance automatisée et continue de populations entières — représente l'une des transformations les plus profondes dans l'exercice du pouvoir, le rôle de la technologie et les limites de l'autonomie individuelle. Ce qui nécessitait autrefois un effort humain important, une priorisation institutionnelle et une sélection délibérée a évolué vers une infrastructure fluide qui génère, agrège et analyse des données comportementales sur des milliards de personnes en temps réel, souvent comme un sous-produit incident de la vie quotidienne.

Cette transformation n'était pas prédéterminée par la technologie seule. Elle est née de l'interaction entre l'expansion bureaucratique, les crises de sécurité répétées, les incitations économiques liées à la monétisation des données et la réduction incessante du coût marginal de la collecte, du stockage et de l'inférence des données. Le résultat n'est pas simplement « plus de surveillance », mais un régime qualitativement différent : il remplace le frottement naturel par une échelle sans frottement, la discrétion humaine par l'automatisation algorithmique, et la suspicion exceptionnelle à l'égard de quelques-uns par l'observation de base de tous.

En son cœur se trouve une métamorphose fondamentale : le profilage est passé d'un **artisanat** — sélectif, gourmand en main-d'œuvre et explicatif — à un **processus industriel** — universel, automatisé et prédictif. Ce qui suit retrace cette transformation, en identifiant les moments où les contraintes se sont érodées et où de nouvelles capacités se sont cristallisées en un système d'inférence continue à l'échelle de la population.

I. Fondations : Le profilage comme pratique sélective et manuelle

Le profilage, dans sa forme la plus élémentaire, consiste en la collecte et l'interprétation systématiques d'informations afin d'inférer des caractéristiques, de prédire des comportements ou d'attribuer des catégories de risque. Ses origines remontent profondément à l'Antiquité.

Les empires anciens menaient des recensements non seulement pour la taxation ou la conscription, mais aussi pour la classification. Les autorités romaines et les administrateurs impériaux chinois triaient les populations par profession, loyauté et statut, produisant des cartes relationnelles précoces capables d'identifier des menaces potentielles. Les institutions religieuses tenaient des registres des naissances, des mariages, des confes-

sions et de la conduite morale, construisant des proto-graphes sociaux qui révélaient des réseaux d'influence et de déviation.

Pourtant, ces systèmes partageaient une contrainte fondamentale : **l'information était coûteuse**. La collecte, la vérification, le stockage et l'interprétation des données exigeaient un travail humain important. Par conséquent, le profilage restait **sélectif, épisodique et limité**. Il se concentrait sur les élites, les dissidents ou les groupes stratégiquement pertinents — et non sur des populations entières.

Cette sélectivité a persisté dans l'Europe moderne naissante, même lorsque les États élargissaient leur appareil de surveillance. Les efforts de renseignement ciblaient les hérétiques, les rivaux politiques, les contrebandiers et les agents étrangers par le biais d'informateurs, de correspondances interceptées et de surveillance physique. Les *cabinets noirs* de la France et d'autres États incarnaient cette approche : des équipes de clercs ouvraient manuellement les lettres, les copiaient et les refermaient pour livraison. Ces opérations étaient intrinsèquement contraintes. Elles se concentraient sur des cibles de haute valeur parce que toute extension plus large était logistiquement impossible.

Même à ce stade, cependant, le pouvoir des **métadonnées** était clairement compris. Les informations sur la communication — expéditeur, destinataire, moment et itinéraire — pouvaient exposer des réseaux et des intentions sans nécessiter l'accès au contenu. Le scandale d'espionnage du bureau de poste britannique de 1844 a mis cela en évidence auprès du public. Le révolutionnaire italien Giuseppe Mazzini, exilé à Londres, soupçonnait que ses lettres étaient ouvertes par les autorités à la demande de puissances étrangères. Lui et ses partisans plaçaient des graines de pavot et des grains de sable à l'intérieur des enveloppes comme marqueurs ; lorsque les lettres arrivaient perturbées, Mazzini incita le député radical Thomas Duncombe à soulever la question au Parlement. Le scandale qui s'ensuivit révéla l'ouverture systématique des lettres sous mandat délivré par le ministre de l'Intérieur Sir James Graham, provoquant l'indignation, des enquêtes parlementaires et l'abolition éventuelle du département secret du bureau de poste. Il marqua l'une des premières paniques modernes sur la vie privée et souligna comment les seules données relationnelles pouvaient démanteler des réseaux d'association.

En réponse, des normes juridiques telles que le « secret de la correspondance » (*Briefgeheimnis*, *secret de la correspondance*) ont émergé. Ces principes limitaient strictement l'utilisation des données de communication à des fins opérationnelles comme la livraison, interdisant leur exploitation secondaire pour la surveillance ou le profilage. L'idée sous-jacente était simple mais profonde :

> Les données générées pour une fonction spécifique ne doivent pas être réutilisées pour construire des profils plus larges d'individus ou de réseaux.

Ce principe résonnerait à travers les siècles — mais finirait par s'éroder sous la pression technologique et institutionnelle.

II. Le Siècle bureaucratique : L'échelle sans automatisation

Le XXe siècle a considérablement élargi le profilage tout en préservant bon nombre de ses contraintes antérieures. Les exigences de la guerre totale nécessitaient une collecte d'informations sans précédent. La censure du courrier, le renseignement d'origine électromagnétique et le décryptage ont étendu la surveillance au-delà des élites vers des populations plus larges. Des institutions telles que la National Security Agency ont institutionnalisé l'interception à grande échelle, tandis que les agences nationales compilaient des dossiers étendus sur les groupes politiques, les radicaux présumés et les réseaux criminels.

Pourtant, le profilage restait **fondamentalement ciblé**. Les écoutes téléphoniques étaient liées à des individus ou des lignes spécifiques. Les dossiers de renseignement étaient élaborés par des analystes humains. Même si le volume augmentait, **l'attention humaine restait le goulot d'étranglement**.

Les premiers systèmes informatiques (années 1950-1970) ont commencé à modifier l'échelle de la tenue des registres. Les gouvernements et les entreprises ont numérisé les listes d'aide sociale, les historiques de crédit et les bases de données criminelles, permettant une récupération et un recoupement plus rapides. Mais ces systèmes fonctionnaient encore sur des **enregistrements discrets**, et non sur des flux continus de comportements.

Dans les années 1970, les préoccupations concernant les « banques de données » centralisées ont suscité des réponses juridiques. La Privacy Act américaine de 1974 et les premières lois européennes sur la protection des données ont introduit les principes de limitation de la finalité, de minimisation des données et de transparence. Ces cadres ont étendu la logique du secret de la correspondance à l'ère numérique.

Cependant, ils reposaient sur une hypothèse cruciale : que la collecte de données était **limitée et épisodique**. Ils réglementaient les enregistrements — et non les flux. Cette hypothèse allait bientôt s'effondrer.

III. Le Point d'inflexion : Des enregistrements aux effluents de données

La rupture décisive se produit à la fin des années 1990 et au début des années 2000 avec l'essor d'Internet — non seulement comme moyen de communication, mais comme infrastructure qui produit continuellement des données.

Les systèmes numériques génèrent des **effluents de données** : des métadonnées créées automatiquement comme sous-produit d'une activité ordinaire. Chaque connexion, requête, clic et mouvement produit des traces qui peuvent être enregistrées, stockées et analysées à un coût négligeable.

Cela marque le tournant décisif :

> Le profilage cesse d'être une activité exercée sur les données pour devenir une infrastructure qui les produit continuellement.

Les fournisseurs d'accès à Internet capturent les journaux de connexion, les requêtes DNS et les informations de routage, révélant des schémas de comportement même sans accès au contenu. Contrairement aux métadonnées postales — éphémères et décentralisées — les métadonnées numériques sont persistantes, centralisées et trivialement consultables.

Au-dessus de cette infrastructure, des plateformes telles que Google et Meta ont transformé le profilage en un modèle économique central. Les moteurs de recherche capturent l'intention ; les réseaux sociaux cartographient les relations ; les écosystèmes mobiles suivent les mouvements. Des traqueurs intégrés étendent la visibilité à de vastes portions du web. Les pixels de suivi de Meta, présents sur environ un tiers des sites web populaires dans le monde, surveillent l'activité bien au-delà de leurs propres plateformes, capturant souvent des signaux sensibles issus de contextes sanitaires, financiers ou politiques.

Une prise de conscience critique émerge dans cet environnement :

> Le contenu devient largement redondant. Dans de nombreux cas, les schémas relationnels ne sont pas simplement des substituts du sens — ils sont plus utiles analytiquement que le contenu lui-même.

Les métadonnées n'indiquent pas simplement qu'une communication a eu lieu ; elles permettent une **reconstruction probabiliste du contenu**. Qui communique avec qui, quand, à quelle fréquence et dans quel contexte plus large peut fortement contraindre ce qui est communiqué. Les informations publiquement disponibles — affiliations partagées, rôles professionnels, positions politiques, liens sociaux — rétrécissent encore l'espace des interprétations plausibles.

Au fil du temps, ces contraintes deviennent prédictives. Les métadonnées ne sont pas seulement descriptives ; elles sont génératives. Elles n'accompagnent pas simplement le contenu — elles peuvent souvent **l'approximer ou l'inférer**, surtout lorsqu'elles sont agrégées à grande échelle.

Les requêtes de recherche révèlent l'intention. La fréquence des communications révèle la force des relations. La co-localisation révèle l'association. À une échelle suffisante, ces signaux convergent vers des modèles comportementaux hautement précis qui rendent souvent l'accès direct au contenu inutile.

Les systèmes d'entreprise optimisent le comportement pour la monétisation ; les systèmes étatiques le contraignent pour le contrôle — mais les deux reposent sur la même machinerie sous-jacente : **la prédiction par inférence comportementale à grande échelle**.

IV. L'Identité sans échappatoire : Ancres persistantes

Une caractéristique déterminante du profilage industriel est l'émergence d'une **identité persistante**.

Les systèmes antérieurs reposaient sur des identifiants mutables — noms, documents, adresses — qui pouvaient être modifiés ou masqués. Les systèmes modernes reconstruisent l'identité à travers des signaux superposés :

- Empreintes digitales des appareils
- Schémas comportementaux
- Graphes sociaux
- Marqueurs biométriques (visages, démarche, voix)

Les images partagées publiquement servent d'ancres durables. Même lorsque les individus changent de comptes ou adoptent des pseudonymes, les systèmes de reconnaissance faciale — particulièrement dans les contextes étatiques ou de renseignement — peuvent reconnecter les identités à travers les ensembles de données. La co-occurrence dans des photos ou des événements partagés renforce encore les relations inférées.

L'implication est profonde :

> L'identité n'est plus quelque chose que l'on déclare, mais quelque chose qui est continuellement inféré.

Cela élimine une grande partie du frottement qui contraignait autrefois la surveillance. L'identification ne dépend plus d'un signal unique ; elle émerge de la redondance à travers de nombreux signaux.

V. La Fusion : Des points de données aux ontologies

L'aboutissement de cette évolution est la **fusion des données** : l'intégration d'ensembles de données disparates dans des systèmes analytiques unifiés.

Des plateformes telles que Palantir Technologies agrègent les dossiers gouvernementaux, les transactions financières, l'activité des médias sociaux, les données de localisation et les métadonnées de communications en modèles cohérents d'individus et de réseaux. Ces systèmes construisent des ontologies dynamiques qui permettent aux analystes d'interroger les relations, de détecter des schémas et de générer des prédictions.

Un exemple concret illustre le changement. Dans l'application des lois sur l'immigration, l'outil ELITE (Enhanced Leads Identification and Targeting for Enforcement) de Palantir peuple des cartes avec des cibles potentielles, en s'appuyant sur les dossiers de visas, les données d'emploi, les métadonnées téléphoniques, les connexions sociales et même les informations d'adresse de Medicaid ou du HHS pour attribuer des « scores de confiance d'adresse » et générer des dossiers. Les agents peuvent identifier des quartiers « riches en cibles » pour des opérations, en signalant des individus non seulement sur la base de preuves directes, mais parce que leur **signature comportementale et relationnelle** ressemble à des cas précédemment identifiés. Une fusion similaire apparaît dans des outils comme ImmigrationOS, qui intègre les historiques de voyage, les données biométriques et les données sociales pour la priorisation.

Le soupçon n'est plus découvert — il est **généré**.

Le profilage ne documente pas seulement la réalité ; il la construit activement en faisant surface des associations probabilistes qui deviennent opérationnellement actionnables.

VI. De l’explication à la prévention anticipée

Le profilage traditionnel était largement rétrospectif. Il cherchait à expliquer les actions passées — qui a commis un crime, qui a organisé un complot, qui constituait une menace.

Le profilage industriel est prédictif et préventif. Il identifie :

- Qui pourrait commettre un crime
- Où un crime pourrait se produire
- Qui pourrait faire défaut, se radicaliser ou dévier

Cette logique est souvent comparée à la vision dépeinte dans *Minority Report*, où des individus sont appréhendés avant de commettre des crimes. Bien que les systèmes contemporains manquent de prévoyance déterministe, la ressemblance structurelle est claire : les outils de police prédictive analysent les données historiques, les appels au 911, les lecteurs de plaques d’immatriculation et les signaux sociaux pour générer des « listes chaudes » ou des scores de risque.

Les systèmes modernes fonctionnent sur la probabilité. Les individus sont signalés non pas parce qu’ils agiront, mais parce qu’ils **ressemblent statistiquement à d’autres qui l’ont fait**.

Le changement est subtil mais profond :

> Les individus ne sont plus jugés principalement sur leurs actions, mais sur leur position dans un paysage probabiliste.

Le soupçon devient structurel — généré continuellement plutôt que déclenché par des événements discrets.

VII. Le Droit à l’ère de l’inférence

Les cadres juridiques tels que le Règlement général sur la protection des données tentent d’imposer des limites par le consentement, la transparence et la minimisation. Pourtant, ils font face à des contraintes structurelles.

La plupart des systèmes juridiques réglementent **les données en tant qu’objet**. Le profilage moderne tire son pouvoir des **relations et des inférences**, qui sont beaucoup plus difficiles à définir, à observer ou à contraindre.

Les défis supplémentaires incluent :

- Les flux continus de données à travers les juridictions
- Les exceptions larges pour la sécurité nationale et les « intérêts légitimes »
- Les systèmes algorithmiques opaques résistants au contrôle

Le résultat est un décalage persistant :

> Les cadres juridiques conçus pour un âge des enregistrements peinent à gouverner un âge de l'inférence prédictive continue.

VIII. L'Asymétrie du pouvoir

Le profilage industriel produit un déséquilibre structurel.

Les individus génèrent des données continuellement par leur participation à la vie moderne. L'évitement est possible mais coûteux et incomplet. Pendant ce temps :

- Les entreprises maintiennent des systèmes opaques protégés par le secret
- Les États accèdent aux données et les intègrent par autorité légale ou partenariats
- La complexité technique obscurcit la responsabilité

Le résultat est une asymétrie claire :

> Les nombreux deviennent lisibles ; les puissants restent comparativement opaques.

IX. L'Internalisation : Le profilage et l'autorégulation du comportement

Au-delà de ses dimensions institutionnelles et technologiques, l'industrialisation du profilage produit une transformation psychologique profonde. La surveillance ne fonctionne plus uniquement comme une force externe ; elle devient internalisée.

Cette dynamique a été anticipée par Michel Foucault dans son analyse du panoptique : un design de prison théorique de Jeremy Bentham dans lequel les détenus, visibles par un observateur central qu'ils ne peuvent pas voir, internalisent la discipline et régulent leur propre comportement sous l'incertitude d'une surveillance constante. Le pouvoir du panoptique réside non pas dans l'observation perpétuelle mais dans l'**anticipation** de celle-ci.

Le profilage industriel étend dramatiquement cette logique. Les individus évoluent dans des environnements où leurs actions peuvent être enregistrées, analysées et interprétées de manière opaque — par des plateformes optimisant l'engagement ou des États évaluant le risque. Le résultat est un virage vers l'**autorégulation**.

Cela se manifeste par :

- L'autocensure dans les publications, les recherches ou les associations
- L'évitement de certains groupes, sujets ou lieux
- L'alignement sur des normes perçues pour minimiser les scores de risque
- La modification du comportement à travers les contextes numériques et physiques

Crucialement, ces adaptations ne nécessitent pas de coercition explicite. Elles naissent de l'anticipation.

> Le contrôle s'exerce non seulement par ce que les systèmes font, mais par ce que les individus évitent de faire.

Les effets s'étendent au-delà des individus. À mesure que les gens s'autocensurent et se trient eux-mêmes, les données générées renforcent les schémas, façonnant les prédictions futures. Le système n'observe pas seulement la réalité — il la remodèle subtilement, créant des boucles de rétroaction qui normalisent la conformité.

X. La Fin de la surveillance sélective

Le profilage a subi une transformation fondamentale :

- De **ciblé** à **universel**
- De **manuel** à **automatisé**
- De **rétrospectif** à **prédictif**
- De **fragmenté** à **intégré**

Les systèmes antérieurs étaient contraints par le frottement — coût, temps, attention humaine. Les systèmes industriels suppriment ces contraintes. La surveillance devient ambiante. L'inclusion devient la norme.

Le principe selon lequel les données ne doivent servir que leur finalité immédiate a cédé la place à un paradigme dans lequel **toutes les données sont potentiellement exploitables**.

XI. Conclusion : Le Prix de la participation

L'arc long du secret postal à la fusion des données numériques révèle un schéma cohérent : chaque expansion technologique augmente la portée du profilage, tandis que les réponses juridiques et sociales tardent. Ce qui distingue le présent est structurel. Le profilage n'est plus une activité dirigée vers des individus spécifiques — c'est une infrastructure au sein de laquelle les individus existent.

La catégorie de « personne d'intérêt » se dissout. Chacun devient sujet à une évaluation continue.

Cette transformation est soutenue non seulement par le pouvoir de l'État, mais aussi par des incitations économiques. Les plateformes qui paraissent gratuites fonctionnent grâce à l'extraction de données comportementales. La phrase *« si vous ne payez pas pour le produit, vous êtes le produit »* capture une intuition — mais sous-estime la réalité.

Ce qui est produit n'est pas l'individu, mais un **modèle prédictif** de l'individu — portable, actionnable et souvent inaccessible à la personne qu'il représente.

Un défi central réside dans l'écart entre la perception et la réalité.

D'abord, les gens sous-estiment **l'impact** de ce qui est connu. Le profilage opère par association. Les relations — passées, faibles ou indirectes — peuvent façonner les résultats.

Une connexion avec quelqu'un qui devient ensuite indésirable peut influencer les opportunités. On est jugé non seulement individuellement, mais relationnellement.

Ensuite, les gens sous-estiment **l'étendue** de ce qui peut être connu. Les systèmes infèrent des attributs sensibles — politiques, religieux, sexuels, économiques — non à partir d'une divulgation explicite, mais à partir de schémas. Ces inférences deviennent opérationnelles quelle que soit leur exactitude.

> Les individus sont évalués non seulement sur ce qu'ils révèlent, mais sur ce qui peut être inféré — et sur ceux à qui ils sont connectés.

La participation à la vie numérique implique donc un échange implicite : la commodité contre la lisibilité. Cet échange n'est ni transparent ni négociable.

Le défi n'est pas d'arrêter la datafication, mais de la contraindre — de restaurer le frottement, d'imposer des limites et d'assurer la responsabilité.

La question centrale est claire :

> L'intervention interviendra-t-elle avant que l'infrastructure du profilage permanent ne devienne trop profondément ancrée pour être contestée de manière significative ?

En l'absence d'une telle intervention, le coût de la participation n'est pas seulement les données — mais l'érosion progressive de la frontière entre être observé, être inféré et, ultimement, être défini.