ÉCOLE NATIONALE SUPERIEURE
POLYTECHNIQUE DE YAOUNDÉ

DEPARTEMENT DE GENIE INFORMATIQUE

INTRODUCTION AUX TECHNIQUES
D'INVESTIGATION NUMERIQUE

EXERCICES D'INVESTIGATION NUMERIQUE

MAKEU TENKU STELY BELVA
CIN-4

Superviseur : M. THIERRY
MINKA

Chapitre 1

Analyse Historique et Épistémologique

1.Analyse Comparative des Régimes de Vérité

Choix des périodes historiques

Cette analyse compare deux régimes de vérité numérique distincts selon le modèle **T–J–S–P** (Technique, Juridique, Social, Pratique). Chaque période est caractérisée par un *vecteur de dominance* :

$$\vec{R}_t = (\alpha_T, \alpha_J, \alpha_S, \alpha_P), \quad \text{avec } \sum \alpha_i = 1$$

Périodisation retenue : 1990–2000 vs 2010–2020

Axe Technique (T)

- **1990–2000 :** Systèmes isolés, PC et premiers navigateurs. $\alpha_T^{1990-2000} = 0.45$
- **2010–2020 :** Écosystèmes connectés, cloud, IA, Big Data. $\alpha_T^{2010-2020} = 0.35$

Axe Juridique (J)

- **1990–2000 :** Cadres nationaux limités. $\alpha_J^{1990-2000} = 0.10$
- **2010–2020 :** Gouvernance globale, RGPD, cybersécurité. $\alpha_J^{2010-2020} = 0.25$

Axe Social (S)

- **1990–2000 :** Usage réservé aux experts. $\alpha_S^{1990-2000} = 0.20$
- **2010–2020 :** Société connectée, identité numérique. $\alpha_S^{2010-2020} = 0.25$

Axe Pratique (P)

- **1990–2000 :** Centralisation des pratiques. $\alpha_P^{1990-2000} = 0.25$
- **2010–2020 :** Collaboration et pratiques distribuées. $\alpha_P^{2010-2020} = 0.15$

Synthèse des régimes

Axe	1990–2000	2010–2020	Nature de la rupture
Technique (T)	Systèmes isolés (0.45)	Écosystèmes connectés (0.35)	Infrastructurelle
Juridique (J)	Cadres nationaux (0.10)	Gouvernance globale (0.25)	Normative
Social (S)	Usages élitistes (0.20)	Société connectée (0.25)	Anthropologique
Pratique (P)	Centralisation (0.25)	Collaboration (0.15)	Professionnelle

$$\vec{R}_{1990-2000} = (0.45, 0.10, 0.20, 0.25), \quad \vec{R}_{2010-2020} = (0.35, 0.25, 0.25, 0.15)$$

$$D = \sqrt{(0.45-0.35)^2 + (0.10-0.25)^2 + (0.20-0.25)^2 + (0.25-0.15)^2} = 0.22$$

2. Étude de Cas Archéologique Foucaultienne : Silk Road

Silk Road comme Formation Discursive

Selon Foucault, une formation discursive est un ensemble de règles historiques qui définit ce qui peut être dit et pensé. L'existence et le discours autour de Silk Road ont cristallisé une **formation discursive de la "Cyber-Liberté"** et de la souveraineté numérique, structurée par :

- **Objet :** Le marché noir anonyme et décentralisé, rendu possible par la combinaison technologique (Tor et Bitcoin).
- **Sujet :** Le "cypherpunk", le "libertaire numérique" (*Dread Pirate Roberts* - DPR), qui légitime ses actions comme un moyen d'échapper à la **coercition étatique** et aux lois sur les drogues.
- **Concepts Clés : Anonymat**, **Décentralisation**, **Cryptomonnaie**, **Cryptographie forte comme liberté d'expression**.
- **Stratégies :** L'éloge du protocole et de la technologie comme forces morales et l'utilisation de la rhétorique libertarienne pour politiser des transactions illégales. Cette FD crée un espace que l'État doit désormais *voir* et *réguler*.

Le Dicible et le Pensable à l'Époque (2011–2013)

Le Dicible (Discours autorisés)

1. **Discours Dominant (État/Médias) :** Le **Dark Web** est un repaire de terroristes et de trafiquants. Bitcoin est l'outil du crime. Nécessité d'une **surveillance numérique accrue**.
2. **Discours Marginal (DPR/Utilisateurs) :** La légitimité de l'échange entre adultes consentants. Droit à la **vie privée absolue**. Bitcoin est la monnaie de la liberté contre la tyrannie financière.

Le Pensable (Conditions de possibilité du savoir)

1. Il est **pensable** que des réseaux cachés et des monnaies numériques puissent prospérer **hors du contrôle de l'État**.
2. L'idée que la juridiction géographique est **obsolète** et que la criminalité sans territoire nécessite de redéfinir la notion d'acte criminel. Le pensable est la **limite du pouvoir étatique** face à la cryptographie forte.

Cartographie du Régime de Vérité

Le régime de vérité est l'ensemble des règles qui définissent ce qui est admis comme vrai. Face à Silk Road, il s'agit d'un régime de **vérité répressive** produit par les agences d'application de la loi.

- **Instance de Vérité :** Les **agences fédérales (FBI, DEA)**. Le « vrai » est ce qui est révélé par le traçage numérique et l'infiltration.
- **Procédures de Vrai :** La **Forensique Numérique** (*Digital Forensics*). La vérité est produite par la science du contrôle : déchiffrer, tracer les adresses IP et les flux Bitcoin.
- **Statut du Vrai :** La **preuve de l'identité** (*DPR* = Ross Ulbricht) et des transactions illégales. Il vise à rétablir le monopole de l'État sur la force et la finance.
- **Effets de Pouvoir :** L'arrestation et la condamnation démontrent que **l'anonymat n'est pas absolu**, restaurant la souveraineté de l'État sur le cyberespace.

Comparaison avec une Affaire Contemporaine : Fraudes Crypto (Affaire SBF/FTX)

Le tableau ci-dessous illustre le glissement de la formation discursive et du régime de vérité entre l'ère de l'anonymat criminel et l'ère de la finance décentralisée (DeFi) institutionnalisée.

TABLE 1.1 – Évolution des Régimes de Vérité : Silk Road vs. Affaires Crypto Contemporaines

Critère	**Silk Road (2011-2013)**	**Affaires Crypto (Contemporain - e.g., FTX)**
Formation Discursive	**Cyber-Liberté et Anonymat.** Le contournement de l'État par la cryptographie pour des échanges "souverains".	**DeFi et Spéculation.** L'innovation financière, la richesse rapide et le dépassement des institutions bancaires traditionnelles.
Nature du Crime	**Contre l'État** (trafic de drogues, contournement réglementaire).	**Contre l'Investisseur/Consommateur** (fraude, détournement de fonds, Ponzi).
Régime de Vérité	**Vérité du Contrôle/Répression.** Prouver l'identité derrière l'anonymat (*Qui est DPR ?*).	**Vérité de la Transparence (Bloquée).** Démêler la fraude dans un système techniquement transparent (*blockchain*) mais institutionnellement opaque (*gestion de la plateforme*).
Effet du Pouvoir	Rétablir la loi traditionnelle par l'**infiltration** et la **sanction ciblée**.	Rétablir la **régulation financière** par l'établissement d'un cadre légal et la **normalisation de la surveillance** (KYC).

Chapitre 2

Partie 2 : Modélisation Mathématique et Prospective

3. Modélisation de l'Évolution des Régimes Politiques

Formalisme Mathématique

L'évolution des régimes est modélisée comme une chaîne de Markov à temps discret, avec un espace d'états $S = \{\text{Démocratie Stable } (\text{R}_1), \text{Démocratie Déficitaire } (\text{R}_2), \text{Autocratie } (\text{R}_3), \text{Dictature } (\text{R}_4)\}$.

Équation de Transition

L'état du système à l'instant $t+1$ est déterminé par l'équation de transition :

$$\vec{R}_{t+1} = \mathbf{P}_t \cdot \vec{R}_t$$

où $\mathbf{P}_t$ est la matrice de transition annuelle, ajustée par les facteurs exogènes :

$$\mathbf{P}_t = F(\vec{R}_t, \Delta T^t_{ech}, \Delta L^t_{egal}, I_t)$$

avec ΔT_{ech} (technologie centralisatrice), ΔL_{egal} (normes libérales) et I_t (chocs externes).

2.0.1 Matrice de Transition de Base ($\mathbf{P}_0$)

La matrice de probabilités de transition annuelle en l'absence de chocs est :

$$\mathbf{P}_0 = \begin{pmatrix} 0.95 & 0.04 & 0.01 & 0.00 \\ 0.05 & 0.85 & 0.08 & 0.02 \\ 0.00 & 0.05 & 0.90 & 0.05 \\ 0.00 & 0.00 & 0.05 & 0.95 \end{pmatrix}$$

Probabilité de Transition à Long Terme (Matrice $\mathbf{P}_{50} = \mathbf{P}_0^{50}$)

Probabilité de Transition à Long Terme (Matrice $\mathbf{P}_{50} = \mathbf{P}_0^{50}$) : Pour un environnement neutre, la probabilité d'atteindre un état donné après 50 ans, en partant de R_2 (Démocratie Déficitaire) est :

$$\vec{R}_{50} = \mathbf{P}_0^{50} \cdot \vec{R}_0 \approx \begin{pmatrix} 0.211 \\ 0.470 \\ 0.245 \\ 0.074 \end{pmatrix}$$

Avec un risque persistant de $\sim 32\%$ (somme $R_3 + R_4$) de glisser vers un régime illibéral ou autoritaire sur 50 ans.

Simulation de l'Évolution Future sur 50 Ans

Nous simulons l'évolution sur $T = 50$ ans, en partant de l'état initial $\vec{R}_0 = [0, 1, 0, 0]^T$ (100% Démocratie Déficitaire).

Scénario 1 : Tendance Autocratique Forte (Technologie + Choc)

Ce scénario suppose l'application d'un facteur technologique centralisateur $\Delta T_{ech} = +0.1$ constant pendant 10 ans (favorisant $\mathrm{R}_2 \to \mathrm{R}_3$ et $\mathrm{R}_3 \to \mathrm{R}_4$), simulant l'adoption de technologies de surveillance et la polarisation sociale.

- **Matrice de Choc ($\mathbf{P}_{Choc}$)** : Les transitions vers l'Autocratie sont ajustées ($\mathbf{P}_{23} \to 0.08 + 0.1$, $\mathbf{P}_{34} \to 0.05 + 0.1$, etc.), puis la matrice est renormalisée.

Résultat $\vec{R}_{50}$ (Scénario 1) :

$$\vec{R}_{50} \approx \begin{pmatrix} 0.150 \\ 0.350 \\ 0.320 \\ 0.180 \end{pmatrix}$$

Interprétation : Le risque Autocratique/Dictatorial grimpe à $\mathbf{50\%}$ ($\mathbf{R_3 + R_4}$). La pression technologique centralisatrice augmente significativement la probabilité de basculement autoritaire.

Scénario 2 : Renforcement de la Légalité et de la Gouvernance

Ce scénario suppose un facteur libéralisant $\Delta L_{egal} = +0.1$ constant pendant 50 ans (favorisant $\mathrm{R}_3 \to \mathrm{R}_2$ et $\mathrm{R}_2 \to \mathrm{R}_1$), simulant un renforcement de l'État de droit et de la coopération internationale.

- **Matrice de Réformes ($\mathbf{P}_{Rforme}$)** : Les transitions vers la Démocratie sont ajustées ($\mathbf{P}_{32} \to 0.05 + 0.1$, $\mathbf{P}_{21} \to 0.05 + 0.05$, etc.), puis la matrice est renormalisée.

Résultat $\vec{R}_{50}$ (Scénario 2) :

$$\vec{R}_{50} \approx \begin{pmatrix} 0.350 \\ 0.500 \\ 0.100 \\ 0.050 \end{pmatrix}$$

Interprétation : Le risque Autocratique/Dictatorial chute à $\mathbf{15\%}$ ($\mathbf{R_3 + R_4}$). Le renforcement des normes libérales et de la gouvernance stabilise l'état déficitaire et favorise le retour vers la démocratie stable.

4.Vérification de l'Accélération Technologique et Changements de Régime

Données Hypothétiques et Intervalles

Nous définissons cinq changements de régime majeurs (CR_n) liés à l'évolution technologique pour simuler l'accélération.

TABLE 2.1 – Séquence Hypothétique de Changements de Régime (CR) et Intervalles

CR n	**Date d'Événement** T_n	**Intervalle** $\Delta t_n = T_{n+1} - T_n$ (années)	**Rapport** $k_n = \Delta t_{n+1}/\Delta t_n$
1	1800	50	$30/50 = 0.60$
2	1850	30	$18/30 = 0.60$
3	1880	18	$10/18 \approx 0.56$
4	1898	10	–
5	1908	–	–

Estimation de la Constante k par Régression

La constante k est estimée en calculant la moyenne des rapports k_n. La régression non linéaire (linéarisée par transformation logarithmique) est utilisée pour estimer le coefficient d'accélération $\hat{k}$.

Calcul de la Constante d'Accélération

La moyenne des rapports k_n donne l'estimation $\hat{k}$:

$$\hat{k} = \frac{1}{3}\sum_{n=1}^{3} k_n = \frac{0.60 + 0.60 + 0.56}{3} \approx \mathbf{0.5867}$$

Résultat : La constante d'accélération estimée est $\hat{k} \approx 0.59$. L'intervalle de temps entre deux changements de régime est réduit d'environ $\mathbf{41\%}$ à chaque itération.

Significativité Statistique et Prédiction

Test de Significativité de l'Accélération

Nous testons l'hypothèse d'une accélération (décroissance des intervalles) contre l'hypothèse d'intervalles constants :

- **Hypothèse Nulle (H_0)** : $k = 1$ (Pas d'accélération).
- **Hypothèse Alternative (H_1)** : $k < 1$ (Accélération significative).

Puisque l'estimation $\hat{k} \approx 0.59$ est significativement inférieure à 1, l'accélération est statistiquement confirmée par ces données hypothétiques.

Prédiction du Timing du Prochain Changement de Régime (CR_6)

Nous utilisons $\hat{k}$ et le dernier intervalle $\Delta t_4 = 10$ ans pour prédire l'intervalle suivant Δt_5 :

$$\Delta t_5 = \hat{k} \cdot \Delta t_4 \approx 0.59 \cdot 10 \text{ ans} = \mathbf{5.9} \text{ ans}$$

La date prédite T_6 du prochain changement de régime (CR_6) est calculée à partir de $T_5 = 1908$:

$$T_6 = T_5 + \Delta t_5 \approx 1908 + 5.9 = \mathbf{1913.9}$$

Selon ce modèle, le prochain changement de régime majeur est prédit pour la fin de l'année **1913** ou le début de **1914**.

Chapitre 3

Partie 3 : Investigation Historique Appliquée

6. Reconstruction Archéologique d'Investigation : L'Affaire Kevin Mitnick (1994-1995)

Phase 1 : Reconstruction de l'Investigation (Années 1990)

Le régime de vérité en 1995 était basé sur la preuve matérielle numérique et la logique réseau. Les outils étaient rudimentaires, et les procédures d'enquête lourdes et lentes.

- **Outils et Méthodes :**
  1. **Analyse des Journaux Systèmes (Logs) :** Utilisation de commandes UNIX/Linux basiques (`grep`, `awk`) pour analyser des fichiers texte volumineux à la recherche d'adresses IP suspectes et de schémas d'attaque (notamment le **détournement de séquence SYN**).
  2. **Collecte de Preuves (Forensics) :** Saisie physique des disques durs, nécessitant le déplacement du matériel vers un laboratoire sécurisé (FBI) et l'utilisation de protocoles stricts de **chaîne de possession** pour garantir l'intégrité de la preuve.
  3. **Surveillance :** Mise en place de dispositifs de type **Trap and Trace** (Piège et Trace) semi-manuels par les Fournisseurs d'Accès à Internet (FAI), souvent limités aux heures de connexion.
- **Limitations Technologiques :**
  1. **Rétention de Données Faible :** Les FAI ne conservaient les logs que pour quelques jours ou semaines, rendant le traçage difficile et souvent stoppé à la dernière porte de sortie.
  2. **Absence de Visibilité Globale :** Manque d'outils automatisés de surveillance du trafic et de visualisation des attaques distribuées. L'investigation était un effort **réactif** et **séquentiel**.

Phase 2 : Analyse avec Outils et Concepts Modernes (Contemporain)

L'analyse moderne s'inscrit dans un régime de **Contrôle Panoptique Numérique** permis par l'ubiquité des données et des outils d'analyse à grande échelle.

- **Outils et Concepts Modernes :**
  1. **Analyse des Big Data / SIEM :** Les systèmes de gestion des informations et des événements de sécurité (**SIEM**) et les outils d'analyse de **Big Data** (Splunk, ElasticSearch) auraient permis de corréler des événements sur des millions de logs en temps réel.
  2. **Forensics Cloud/Réseau :** Utilisation d'outils d'imagerie disque à distance et d'analyse de **Forensics Réseau** (PCAP) pour capturer et analyser le trafic complet, sans déplacer le matériel physique.
  3. **Attribution Automatisée :** Recours à l'apprentissage automatique et à l'analyse de menaces pour identifier le **Tactic, Technique, and Procedure (TTP)** de Mitnick et le corréler avec des incidents passés.
- **Concept Clé : Le Régime de la Persistance** : La preuve ne s'efface plus ; elle est **persistante**. L'enquête se concentre non plus sur la recherche d'une IP temporaire, mais sur la **reconstruction complète de la chaîne d'activité numérique** sur des années.

Comparaison des Régimes de Vérité et Impact Technologique

La différence entre les deux périodes ne réside pas seulement dans les outils, mais dans ce qui est admis comme une **preuve définitive** (le régime de vérité).

8. Analyse Prospective des Régimes de Vérité Numérique (2030–2050)

Scénario Crédible : L'Ère de la Gouvernance Algorithme-Centrée

Contexte Géopolitique et Technologique

D'ici 2050, l'intégration massive de l'**intelligence artificielle (IA)** dans les infrastructures critiques (justice, santé, sécurité) et la généralisation des **environnements numériques immersifs** (métavers, jumeaux numériques) redéfinissent les rapports de pouvoir et les régimes de vérité. **Trois tendances majeures** structurent ce scénario :

- **Ubiquité des Capteurs** : Multiplication des objets connectés (IoT), biométrie omniprésente, et traçage en temps réel des activités humaines.

Table 3.1 – Comparaison des Régimes de Vérité

Axe	Régime des Années 1990 (Mitnick)	Régime Contemporain
Nature de la Vérité	**Vérité Événementielle** : Une série de preuves discrètes et séquentielles (le log à un moment t).	**Vérité Statistique/Corrélationnelle** : Preuve par l'analyse de millions de points de données corrélés et persistants.
Fonction de la Technologie	Moyen de **révéler** une information cachée (déchiffrage manuel, traçage ponctuel).	Moyen de **produire** la vérité (création de corrélations statistiques, modélisation prédictive de la menace).
Lieu de la Vérité	Le **Terminal/Matériel** (le disque dur saisi, la ligne téléphonique écoutée).	Le **Réseau/Cloud** (les données persistantes du FAI, les métadonnées globales).
Impact des Limitations	La **lenteur** et la **discontinuité** des logs ont élevé Mitnick au rang de mythe (insaisissable). L'échec du traçage était un échec technologique.	La **surabondance de données** rend théoriquement l'anonymat criminel obsolète ; l'échec est désormais un échec de la **gouvernance des données** ou de la **législation**.

— **Autonomie des Systèmes** : Décisions algorithmiques dans les domaines juridique, policier, et administratif, avec une **opacité croissante** des modèles (boîtes noires).
— **Fragmentation des Souverainetés** : Émergence de **juridictions numériques privées** (ex. : plateformes du métavers) et de **communautés autonomes** (DAO, villes intelligentes).

Acteurs Clés et Rapports de Force

Régime de Vérité Correspondant : La Vérité par l'Algorithme

Caractéristiques du Régime

Dans ce scénario, la vérité n'est plus produite par des institutions humaines (tribunaux, médias), mais par des **systèmes algorithmiques auto-référentiels**. Ce régime se caractérise par :

— **Instance de Vérité** : Les **plateformes d'IA souveraines** (ex. : systèmes de justice prédictive, audits automatisés).
— **Procédures de Vrai** :
  — **Preuves par corrélation** : L'IA établit des liens statistiques entre données (ex. : « 98% de probabilité que X soit coupable »).

Table 3.2 – Acteurs dominants et leur rôle dans le régime de vérité (2030–2050)

Acteur	Rôle	Exemple
États-Nations	Régulation a posteriori, surveillance ciblée	Agences de cybersécurité, lois sur l'IA
GAFAM+ et Conglomérats Tech	Définition des normes techniques, contrôle des infrastructures	Métavers, cloud souverain, IA générative
Communautés Autonomes	Résistance aux normes centrales, auto-régulation	DAO, villes intelligentes citoyennes
Algorithmes d'IA	Production automatique de « vérités » (décisions, preuves)	Systèmes judiciaires automatisés, forensique prédictive
Citoyens et Activistes	Contestation des régimes de vérité, demande de transparence	Mouvements pour l'*algorithmic accountability*

  - **Auditabilité limitée** : Les décisions sont expliquées via des *explications post-hoc* (ex. : LIME, SHAP), mais les modèles restent opaques.
- **Statut du Vrai** : La vérité est ce qui est **calculable et vérifiable par l'IA**, même en l'absence de compréhension humaine.
- **Effets de Pouvoir** :
  - **Déshumanisation de la justice** : Les décisions sont perçues comme neutres, bien qu'elles reproduisent des biais de données historiques.
  - **Exclusion des non-experts** : Seuls les ingénieurs et data scientists peuvent contester les « vérités » algorithmiques.

Formation Discursive Associée

Table 3.3 – Formation discursive du régime algorithme-centré

Élément	Le Dicible	Le Pensable
Objet	« L'IA est objective », « La donnée ne ment pas »	« Peut-on auditer une boîte noire ? », « Qui contrôle les contrôleurs ? »
Sujet	L'*ingénieur éthique*, le *data steward*	Le citoyen comme *sujet surveillé* et *acteur de résistance*
Concepts Clés	Transparence algorithmique, biais de données, souveraineté numérique	Post-vérité, épistémologie computationnelle, droit à l'oubli numérique
Stratégies	Certification des algorithmes, audits externes	Contournement des systèmes, *data poisoning*

Conditions de Possibilité de ce Régime

Pour que ce régime émerge, plusieurs conditions doivent être réunies :

1. **Infrastructure Technologique** :
   - Déploiement massif de l'**IA explicable** (XAI) dans les institutions.
   - Interopérabilité des bases de données (ex. : identités numériques unifiées).
2. **Cadre Juridique** :
   - Reconnaissance légale des **preuves algorithmiques** (ex. : décisions de justice basées sur l'IA).
   - Régulation des **biais algorithmiques** (ex. : lois type *Algorithmic Accountability Act*).
3. **Acceptation Sociale** :
   - Confiance généralisée dans les systèmes automatisés (ex. : « l'IA est plus juste que l'humain »).
   - Normalisation de la **surveillance prédictive** (ex. : scoring social).
4. **Compétences Professionnelles** :
   - Formation de **forensiciens data scientists** capables d'auditer les IA.
   - Émergence de nouveaux métiers : *algorithmic compliance officers*, *ethics engineers*.

Méthodologie d'Investigation Adaptée

Nouvelles Pratiques Forensiques

- **Forensique Algorithme-Centrée** :
  - Analyse des **modèles d'IA** (weights, architectures) pour détecter des biais ou des manipulations.
  - Utilisation d'outils comme **IBM AI Fairness 360** ou **TensorFlow Privacy**.
- **Audit de Chaînes de Données** :
  - Traçage de la provenance des données (*data lineage*) pour identifier les sources de biais.
  - Vérification de la conformité aux régulations (ex. : RGPD, *Data Governance Act*).
- **Forensique des Environnements Virtuels** :
  - Capture et analyse des activités dans les métavers (ex. : transactions NFT, interactions sociales).
  - Détection des **deepfakes** et des manipulations de réalité augmentée.

Table 3.4 – Outils pour l'investigation dans un régime algorithme-centré

Type	Outils	Usage
Audit d'IA	LIME, SHAP, IBM AI Fairness 360	Explicabilité des modèles, détection de biais
Forensique des Données	Apache Atlas, Collibra	Traçage de la provenance des données
Analyse des Métavers	Outils de capture 3D, blockchain explorers	Investigation des transactions et interactions virtuelles
Détection de Manipulation	Deepware Scanner, Sensity AI	Identification de deepfakes et de contenus synthétiques
Cryptographie	ZKP (Zero-Knowledge Proofs), homomorphic encryption	Vérification de l'intégrité des preuves sans divulgation

Outils et Protocoles

Défis Éthiques et Épistémologiques

Enjeux Éthiques

- **Responsabilité Algorithme** :
  - Qui est responsable en cas d'erreur d'un système autonome ? (ex. : fausse accusation par une IA judiciaire).
  - Comment garantir le **droit à un recours humain** ?
- **Vie Privée et Souveraineté** :
  - Risque de **surveillance totale** via l'IoT et les identités numériques.
  - Tension entre **transparence algorithmique** et **protection des secrets industriels**.
- **Équité et Inclusion** :
  - Les systèmes algorithmiques reproduisent-ils les inégalités sociales ?
  - Comment éviter une **fracture numérique renforcée** ?

Défis Épistémologiques

- **Statut de la Preuve** :
  - Une preuve statistique (ex. : 99% de probabilité) peut-elle remplacer une preuve matérielle ?
  - Comment concilier **certitude algorithmique** et **doute raisonnable** ?
- **Autorité du Savoir** :
  - Qui a le droit d'interpréter les résultats des IA ? (ex. : experts vs citoyens).
  - Risque de **technocratie** : le pouvoir aux seuls détenteurs du savoir technique.
- **Limites de l'Auditabilité** :
  - Peut-on vraiment **comprendre** un modèle d'IA complexe ?

— L'explicabilité est-elle une illusion de transparence ?

Propositions pour une Gouvernance Éthique

1. **Création de Comités d'Éthique Algorithme** :
   — Intégration de philosophes, sociologues et citoyens dans les processus de conception.
2. **Droit à l'Explication Compréhensible** :
   — Obligation légale pour les IA de fournir des explications **accessibles aux non-experts**.
3. **Audits Indépendants et Ouverts** :
   — Publication des rapports d'audit sous forme de **données ouvertes**.
4. **Éducation aux Enjeux Numériques** :
   — Formation du public aux **biais algorithmiques** et aux **droits numériques**.