OBSERVIA

Observatoire des Mutations Algorithmiques

CARTOGRAPHIE EXHAUSTIVE

Fiches Techniques • Matrices d'Impact • Chronologies
Grilles d'Audit • Analyses Sectorielles

Kit méthodologique complet pour l'analyse approfondie
des transformations algorithmiques de l'information

VERSION EXHAUSTIVE — 80+ PAGES

Janvier 2026

Table des matières

Section 1 — Introduction méthodologique

Cette cartographie exhaustive constitue l'outil de référence pour l'analyse approfondie des mutations algorithmiques. Elle rassemble l'ensemble des instruments nécessaires à une investigation rigoureuse : fiches techniques détaillées, matrices d'impact multidimensionnelles, chronologies annotées, grilles d'audit complètes, et analyses sectorielles.

Le document est structuré pour permettre à la fois une lecture linéaire (découverte progressive des outils) et une consultation ponctuelle (accès direct à un instrument spécifique). Chaque outil est présenté avec son mode d'emploi, ses critères d'évaluation, et des exemples concrets tirés de l'analyse des trois systèmes paradigmatiques : Google PageRank (paradigme citationnel), Facebook News Feed (paradigme comportemental), et GPT-4 (paradigme génératif).

1.1 Principes directeurs

L'analyse des mutations algorithmiques repose sur cinq principes méthodologiques fondamentaux qui guident l'ensemble de la cartographie :

- EXHAUSTIVITÉ : Couvrir toutes les dimensions pertinentes (technique, économique, sociale, démocratique, éthique) sans en privilégier aucune a priori.
- TRAÇABILITÉ : Documenter les sources, les méthodes de collecte, et les critères d'évaluation pour permettre la vérification et la reproductibilité.
- COMPARABILITÉ : Utiliser des cadres d'analyse standardisés permettant la comparaison systématique entre systèmes et entre périodes.
- CRITICITÉ : Maintenir une posture analytique qui interroge les présupposés, les non-dits, et les effets de pouvoir des systèmes étudiés.
- OPÉRATIONNALITÉ : Produire des outils directement utilisables par les praticiens (auditeurs, régulateurs, chercheurs, décideurs).

1.2 Structure du document

Le document s'organise en huit sections principales, chacune correspondant à un type d'outil ou d'analyse :

1. Fiches techniques exhaustives — Documentation approfondie de chaque système (architecture, données, performances, risques)
2. Matrices d'impact multidimensionnelles — Évaluation systématique des effets sur 10+ dimensions
3. Chronologie annotée — 60+ événements clés avec classification et analyse
4. Grilles d'audit — Checklists opérationnelles pour l'évaluation de conformité
5. Analyses comparatives — Synthèses transversales entre paradigmes
6. Indicateurs de risques — Tableaux de bord et radars de vulnérabilité
7. Études sectorielles — Impacts par domaine (médias, santé, éducation, emploi)
8. Recommandations — Actions pour organisations, régulateurs, utilisateurs

Section 2 — Fiches techniques exhaustives

Les fiches techniques constituent le socle documentaire de la cartographie. Elles rassemblent l'ensemble des informations disponibles sur chaque système analysé, structurées en sections standardisées permettant la comparaison. Chaque fiche couvre dix dimensions : identification, fondements théoriques, architecture technique, infrastructure, signaux/features, évolutions majeures, métriques, biais et limites, impacts économiques, gouvernance.

Les informations proviennent de sources multiples : publications scientifiques, documentation technique officielle, rapports d'audit, articles journalistiques, leaks et témoignages internes, études académiques indépendantes. Le niveau de certitude varie selon les sources ; les estimations sont explicitement signalées.

2.1 Fiche technique — Google PageRank

Système fondateur du paradigme citationnel. Analyse couvrant la période 1998-2012, de la création à Stanford jusqu'au basculement vers les signaux comportementaux et l'ère Knowledge Graph.

FICHE TECHNIQUE EXHAUSTIVE GOOGLE PAGERANK — Système de classement par analyse de liens Période d'analyse : 1998-2012 | Version : Évolution complète	
1. IDENTIFICATION ET CONTEXTE	
Nom officiel	PageRank Algorithm (PR)
Inventeurs	Larry Page, Sergey Brin (Stanford University, 1996-1998)
Publication fondatrice	'The Anatomy of a Large-Scale Hypertextual Web Search Engine' (WWW7, 1998)
Brevet	US Patent 6,285,999 (déposé 1998, expiré 2019)
Opérateur	Google LLC (Alphabet Inc. depuis 2015)
Déploiement initial	Septembre 1998 (google.stanford.edu → google.com)
Périmètre géographique	Mondial — 90%+ parts de marché search (2004-2012)
Utilisateurs affectés	~1 milliard d'utilisateurs quotidiens (2010)
Secteurs impactés	Search, Publicité digitale, E-commerce, Médias, Tourisme, Finance
2. FONDEMENTS THÉORIQUES ET ÉPISTÉMOLOGIQUES	
Paradigme épistémologique	Citationnel — La pertinence émerge de la reconnaissance collective encodée dans les hyperliens
Inspiration scientifique	Science Citation Index (Eugene Garfield, 1964), Impact Factor, Analyse bibliométrique
Hypothèse fondatrice	Un lien hypertexte = un vote de confiance. La qualité d'une page est fonction de la qualité des pages qui la citent (définition récursive)
Modèle conceptuel	Random Surfer Model — Surfeur aléatoire naviguant sur le web en suivant les liens avec probabilité uniforme

Théorie sous-jacente	Théorie des graphes, Chaînes de Markov, Processus stochastiques, Algèbre linéaire (vecteurs propres)
Présupposé normatif	Le web forme une communauté épistémique dont les jugements agrégés révèlent la qualité objective des contenus
3. ARCHITECTURE TECHNIQUE DÉTAILLÉE	
Type d'algorithme	Calcul itératif de vecteur propre sur matrice stochastique (power iteration)
Structure de données	Graphe orienté G = (V, E) où V = pages web, E = hyperliens
Matrice d'adjacence	M[i,j] = 1/Outdegree(j) si lien de j→i, sinon 0. Matrice colonne-stochastique.
Formule mathématique	PR(p) = (1-d)/N + d × Σ[PR(q)/L(q)] pour tout q pointant vers p
Notation matricielle	π = d × M × π + (1-d)/N × e, où π = vecteur PageRank, e = vecteur unitaire
Facteur d'amortissement (d)	d ≈ 0.85 (probabilité de suivre un lien vs téléportation aléatoire)
Justification du damping	Résout le problème des rank sinks (pages sans liens sortants) et garantit l'ergodicité de la chaîne
Convergence	O(log N) itérations pour convergence ε < 10^{-6} (typiquement 50-100 itérations)
Complexité spatiale	O(N) pour le vecteur PR + O(E) pour la matrice sparse
Complexité temporelle	O(k × E) par itération où k = nombre d'itérations jusqu'à convergence
Distribution parallèle	MapReduce (2004+) : Map = calcul contributions, Reduce = agrégation par page
4. INFRASTRUCTURE ET DONNÉES	
Crawler	Googlebot — Exploration BFS/DFS du graphe web, respect robots.txt, scheduling intelligent
Fréquence de crawl	Variable : sites majeurs = quotidien ; long tail = mensuel/trimestriel
Index primaire	~60 trillion URLs connues, ~30 billion pages indexées (2012)
Graphe de liens	~1 trillion d'arêtes (liens) entre pages indexées
Stockage	Bigtable (2004+), GFS, Colossus. Plusieurs centaines de PB.
Calcul PageRank	Batch processing mensuel initialement, puis quasi-continu avec infrastructure distribuée
Datacenters	~15 datacenters mondiaux (2012), >1 million de serveurs
Latence query	< 200ms end-to-end (objectif < 100ms)
5. SIGNAUX DE CLASSEMENT COMPLÉMENTAIRES	
Nombre total de signaux	>500 signaux combinés (estimation 2012), pondérés par ML
Signaux on-page	Title tag, H1-H6, keyword density, TF-IDF, proximity, anchor text, alt text, meta description
Signaux de liens	PageRank, anchor text, contexte sémantique du lien, diversité des domaines référents, velocity des liens
Signaux techniques	Vitesse de chargement, mobile-friendliness (2010+), HTTPS (2014+), structured data
Signaux utilisateurs	CTR, pogo-sticking, dwell time, search refinements (introduits progressivement 2009+)
Signaux de fraîcheur	Query Deserves Freshness (QDF) — boost pour actualités, événements récents
Signaux locaux	IP géolocation, Google My Business, reviews, NAP consistency

Filtres anti-spam	Détection link farms, cloaking, keyword stuffing, doorway pages, hidden text
6. ÉVOLUTION ET MISES À JOUR MAJEURES	
Florida (2003)	Première major update. Pénalisation keyword stuffing. Séisme dans l'industrie SEO.
Big Daddy (2005)	Refonte infrastructure crawl/index. Meilleure gestion canonicalization.
Universal Search (2007)	Intégration résultats verticaux (images, vidéos, news, maps) dans SERP principale.
Caffeine (2010)	Nouvelle architecture d'indexation. Index 50% plus frais, crawl continu.
Panda (2011)	Algorithme qualité contenu. Pénalise thin content, content farms, duplicate. ~12% requêtes affectées.
Penguin (2012)	Algorithme qualité liens. Pénalise link schemes, anchor text sur-optimisé. ~3% requêtes affectées.
Knowledge Graph (2012)	Passage de 'strings to things'. Base de connaissances structurée. Réponses directes.
Fréquence updates	~500-600 changements/an (2012), soit ~1.5/jour en moyenne
7. MÉTRIQUES ET ÉVALUATION DES PERFORMANCES	
Métriques IR classiques	Precision@k, Recall@k, F1-score, Mean Average Precision (MAP), NDCG
NDCG@10 (estimé)	0.75-0.85 selon type de requête (navigational > informational > transactional)
Métriques utilisateurs	Click-through rate (CTR), Time to first click, Pogo-stick rate, Query reformulation rate
Search Quality Raters	>10,000 évaluateurs humains, Search Quality Evaluator Guidelines (160+ pages)
Critères E-A-T	Expertise, Authoritativeness, Trustworthiness — cadre d'évaluation qualité (formalisé 2014)
Benchmarks externes	TREC Web Track, CLEF, NTCIR — participation et domination
Part de marché	~92% mondial (2012), >95% Europe, ~67% US (concurrence Bing)
Volume queries	~5.5 milliards requêtes/jour (2012), ~2 trillion/an
8. BIAIS DOCUMENTÉS ET LIMITES STRUCTURELLES	
Effet Matthieu	Rich-get-richer : pages établies accumulent plus de liens, renforçant leur domination. Auto-amplification.
Biais linguistique	Sur-représentation anglophone (~55% du web indexé). Sous-indexation langues à faibles ressources.
Biais de fraîcheur	Contenu récent sous-évalué (temps d'accumulation de liens). Compensé partiellement par QDF.
Biais de popularité	Corrélation popularité/qualité présupposée mais non garantie. Contenus de niche défavorisés.
Manipulabilité	Link building artificiel, PBN (Private Blog Networks), link wheels, guest posting spam.
Opacité relative	Principe général public, mais pondérations et signaux spécifiques non divulgués.
Inadaptation temps réel	Architecture batch inadaptée aux contenus éphémères (tweets, breaking news).
Inadaptation personnalisation	Classement universel inadapté à la diversité des besoins individuels.

9. IMPACTS ÉCONOMIQUES DOCUMENTÉS	
CA Google (2012)	$50.2 milliards, dont ~95% issus de la publicité search (AdWords)
Marché SEO	~$65 milliards (2012), industrie entièrement créée par l'existence de PageRank
Dépendance éditeurs	30-70% du trafic des sites médias issu de Google Search
Google Tax debate	Controverses sur captation de valeur des contenus (actualités, comparateurs)
Destruction créatrice	Disruption annuaires (Pages Jaunes -80% valeur), agences de voyage, médias classifiés
Barrières à l'entrée	Coût index complet : >$1 milliard. Effet réseau data. Monopole naturel.
10. GOUVERNANCE ET RÉGULATION	
Contrôle interne	Search Quality Team (~1,000 personnes), Core Ranking Team, spam team
Raters guidelines	Document public (depuis 2013), mais application opaque
Webmaster Tools	Interface diagnostic pour webmasters, rapports de liens, erreurs crawl
Procédures de recours	Demandes de réexamen (manual actions), mais pas de recours formalisé utilisateurs
Enquêtes antitrust	FTC (2013 - classée), Commission Européenne (2010+ - sanctions 2017-2019)
Griefs concurrentiels	Self-preferencing (Google Shopping), scraping contenus, abus position dominante
Transparence	Principes généraux publics ; détails algorithme = secret industriel

2.2 Fiche technique — Facebook News Feed

Système emblématique du paradigme comportemental. Analyse couvrant la période 2006-2022, d'EdgeRank au machine learning distribué, incluant les crises Cambridge Analytica et Facebook Files.

FICHE TECHNIQUE EXHAUSTIVE FACEBOOK NEWS FEED — Système de recommandation personnalisée Période d'analyse : 2006-2022 | De EdgeRank au ML distribué	
1. IDENTIFICATION ET CONTEXTE	
Nom du système	Facebook News Feed Ranking Algorithm
Première version	EdgeRank (2006-2011) — formule déterministe Affinity × Weight × Decay
Version ML	Machine Learning-based ranking (2011+), deep learning (2015+)
Opérateur	Meta Platforms, Inc. (Facebook, Inc. jusqu'en 2021)
Utilisateurs	2.91 milliards MAU (2021), 1.93 milliards DAU
Contenus traités	>100 milliards de posts/jour éligibles au ranking
Prédictions/jour	~10^15 prédictions de scoring par jour (estimation)
Équipe dédiée	~1,000 ingénieurs Feed Ranking + Integrity (~15,000 content moderation)
2. FONDEMENTS THÉORIQUES ET PARADIGME	
Paradigme épistémologique	Comportemental — La pertinence = probabilité d'engagement prédite à partir des comportements passés
Théorie implicite	Behaviorisme computationnel : préférences révélées par actions observables (clics, likes, comments)
Hypothèse centrale	Le comportement passé prédit le comportement futur. Les utilisateurs similaires aiment des contenus similaires.
Fonction objectif initiale	Maximiser l'engagement (temps passé, interactions) comme proxy de la valeur utilisateur
Évolution fonction objectif	2018+ : Meaningful Social Interactions (MSI) — prioriser interactions 'significatives' (commentaires, partages)
Tension fondamentale	Optimisation court-terme (engagement) vs bien-être long-terme (non mesurable)
3. ARCHITECTURE EDGERANK (2006-2011)	
Formule	Score = Σ(edges) Affinity(u,e) × Weight(e) × Decay(t)
Affinity (Affinité)	Score relationnel user-creator basé sur historique d'interactions (likes, comments, views, messages)
Weight (Poids)	Pondération par type d'interaction : Comment > Like > Click > View. Photos/Videos > Text.
Decay (Décroissance)	Fonction temporelle décroissante : contenus récents favorisés. Demi-vie ~heures.
Limites EdgeRank	Règles fixes, pas d'apprentissage, ne scale pas avec complexité croissante
4. ARCHITECTURE MACHINE LEARNING (2011-2022)	
Framework	Multi-stage ranking : Candidate Generation → Ranking → Re-ranking → Delivery

Stage 1: Inventory	Collecte des ~1,500 posts éligibles (amis, pages, groupes, ads)
Stage 2: Signals	Extraction features : >10,000 signaux par (user, post) pair
Stage 3: Predictions	Multi-task learning : P(like), P(comment), P(share), P(hide), P(report), etc.
Stage 4: Relevance Score	Combinaison pondérée des probabilités en score final
Stage 5: Contextual	Ajustements contextuels : diversité, fraîcheur, quotas par type
Modèles utilisés	Gradient Boosted Decision Trees (2011), Deep Neural Networks (2015+), Transformers (2019+)
Architecture DNN	Wide & Deep learning, embeddings utilisateur/contenu, attention mechanisms
Features utilisateur	Démographiques, comportementales, sociales, temporelles, device, géolocation
Features contenu	Type, créateur, engagement historique, texte (NLP), image (CV), vidéo features
Features contextuelles	Heure, jour, session history, device state, connection type
Training data	Billions d'interactions/jour, modèles re-trained quotidiennement
Serving infrastructure	Inference temps réel <100ms, GPU clusters, caching agressif
5. PIVOT MSI (2018) — MEANINGFUL SOCIAL INTERACTIONS	
Annonce	11 janvier 2018 (Mark Zuckerberg), suite aux critiques sur bien-être et désinformation
Principe	Prioriser les interactions 'actives' (comments, shares) sur les interactions 'passives' (likes, views)
Changement de scoring	Augmentation poids des contenus amis/famille, réduction reach organique pages/médias
Impact publishers	-50% reach organique médias, accélération déclin du trafic referral
Impact engagement	-5% time spent (court terme), stabilisation qualité perçue
Critiques	MSI favorise aussi contenus clivants (génèrent plus de commentaires)
6. SYSTÈMES D'INTÉGRITÉ ET MODÉRATION	
Architecture Integrity	Classifiers ML parallèles au ranking : hate speech, violence, nudity, misinfo, spam
Modèles de détection	BERT-based pour texte, ResNet pour images, multimodal pour memes
Fact-checking	Partenariat 80+ fact-checkers (IFCN), labels informatifs, réduction distribution
Réduction reach misinfo	-80% distribution pour contenus marqués faux par fact-checkers
Modération humaine	~15,000 modérateurs (outsourcés), review des contenus escaladés
Oversight Board	Créé 2020, 20 membres indépendants, décisions contraignantes sur cas difficiles
Efficacité déclarée	Proactive detection rate : 96%+ hate speech, 99%+ spam (chiffres Facebook)
Limites documentées	Langues non-anglaises sous-couvertes, contexte culturel mal capté, faux positifs
7. DONNÉES COLLECTÉES ET VIE PRIVÉE	
Données on-platform	Profil, posts, likes, comments, shares, messages, groupes, events, photos, videos, reactions
Données comportementales	Scrolling, dwell time, hovers, clicks, play/pause, mute, expand, device orientation

Données off-platform	Facebook Pixel, SDK mobile, Like button, login Facebook (jusqu'en 2018)
Cambridge Analytica	Mars 2018 : révélation collecte données 87M users via app tierce. Scandale majeur.
Conséquences CA	Amende FTC $5B, restriction APIs, audit vie privée, Congressional hearings
Post-RGPD	Consentement explicite EU, data portability, limitation tracking, Privacy Checkup
iOS 14.5 (2021)	App Tracking Transparency : -15% revenus ads estimé, dégradation targeting
8. MÉTRIQUES ET OPTIMISATION	
North Star Metric	DAU/MAU ratio (stickiness), temps passé par session, sessions/jour
Engagement metrics	Likes, comments, shares, clicks, video views, reactions (6 types depuis 2016)
Quality metrics	User surveys (1-5 satisfaction), 'worth it' score, would you miss it?
Integrity metrics	Prevalence of violating content, actioned rate, user reports
Revenue metrics	ARPU ($40/user US, $10 global), ad impressions, CTR, CPM
A/B testing	~10,000 experiments simultanés, allocation utilisateurs, significance testing
Long-term holdouts	Cohortes contrôle sur plusieurs mois pour effets long-terme
9. BIAIS DOCUMENTÉS ET CONTROVERSES	
Amplification émotionnelle	Contenus générant indignation/colère sur-distribués (5x engagement). Confirmé par études internes.
Filter bubbles	Homophilie algorithmique : exposition réduite aux perspectives différentes. Débattu dans la recherche.
Désinformation	Diffusion virale de fausses infos (élections 2016, COVID). Mesures tardives.
Santé mentale ados	Études internes (2019) : Instagram aggrave body image issues chez 32% des teen girls.
Facebook Files (2021)	Fuite documents internes (WSJ). Révèle arbitrages conscients engagement vs sécurité.
XCheck program	Système VIP : 5.8M utilisateurs exemptés des règles standards (politiques, célébrités).
Ethnic violence	Myanmar, Éthiopie : rôle de Facebook dans incitation à la violence. Aveux tardifs.
Biais politiques	Accusations des deux côtés (biais libéral / amplification extrême-droite). Audits contradictoires.
10. IMPACTS ÉCONOMIQUES	
Revenus Meta (2021)	$117.9 milliards (97% publicité)
Valorisation	$900B+ peak (2021), $270B post-iOS14/metaverse pivot (2022)
Marché social ads	~25% du marché pub digitale mondiale
Impact médias	Dépendance trafic (30-50%), pivot to video (2016-2018), effondrement reach
Créator economy	Création d'un marché influenceurs, monétisation via brand deals
Emplois modération	~15,000 emplois précaires, trauma documenté, sous-traitance controversée

2.3 Fiche technique — GPT-4 / ChatGPT

Système inaugural du paradigme génératif. Analyse couvrant la période 2020-2024, de GPT-3 à GPT-4 Turbo, incluant l'architecture Transformer, le pipeline RLHF, et les enjeux d'alignement.

FICHE TECHNIQUE EXHAUSTIVE GPT-4 / CHATGPT — Grand Modèle de Langage Génératif Période d'analyse : 2020-2024 | De GPT-3 à GPT-4 Turbo	
1. IDENTIFICATION	
Nom	GPT-4 (Generative Pre-trained Transformer 4)
Interface grand public	ChatGPT (chat.openai.com) — lancé 30 novembre 2022
Opérateur	OpenAI LP (capped-profit), partenariat Microsoft ($13B)
Chronologie	GPT-1 (2018) → GPT-2 (2019) → GPT-3 (2020) → ChatGPT (2022) → GPT-4 (2023) → GPT-4 Turbo (2023)
Utilisateurs ChatGPT	100M+ utilisateurs hebdomadaires (2023), 1.8B visites/mois
API developers	2M+ développeurs utilisant l'API (2023)
Intégrations	Microsoft Copilot, Bing Chat, GitHub Copilot, Office 365, Azure OpenAI
Valorisation OpenAI	$86B+ (2024), levées $11B+ total
2. FONDEMENTS THÉORIQUES ET PARADIGME	
Paradigme	Génératif — Le système produit du texte nouveau plutôt que de retrouver de l'existant
Épistémologie implicite	Cohérentisme : la vérité = cohérence interne du discours (vs correspondance au réel)
Hypothèse scaling	'Scaling laws' : performances ∝ log(params) × log(data) × log(compute)
Capacités émergentes	Phénomènes non prédits apparaissant au-delà de certains seuils de taille
Théorie controversée	'Stochastic parrots' (Bender et al.) vs 'emergent reasoning' — débat ouvert
Alignement	Problème de faire correspondre les objectifs du modèle aux intentions humaines
3. ARCHITECTURE TRANSFORMER	
Base	Transformer decoder-only (Vaswani et al., 2017, modifié)
Mécanisme clé	Self-attention : Attention(Q,K,V) = softmax(QK^T/√d_k)V
Multi-head attention	h=96 têtes parallèles (GPT-3), chacune capture différents patterns
Positional encoding	RoPE (Rotary Position Embedding) pour séquences longues
Feed-forward	FFN(x) = GELU(xW_1)W_2, dimension intermédiaire 4×d_model
Layer normalization	Pre-LN (avant chaque sous-couche) pour stabilité training
Residual connections	x + Sublayer(x) à chaque couche pour gradient flow
Params GPT-3	175 milliards paramètres, 96 layers, d_model=12288
Params GPT-4 (estimé)	~1.7 trillion params (MoE architecture, 8 experts × 220B)
Context window	GPT-4: 8K/32K tokens, GPT-4 Turbo: 128K tokens
Tokenizer	BPE (Byte Pair Encoding), vocabulaire ~100K tokens

MoE (Mixture of Experts)	Routage sparse : seuls 2 experts activés par token (efficacité compute)
4. PIPELINE D'ENTRAÎNEMENT	
Phase 1: Pretraining	Prédiction next-token sur corpus massif. Objectif : minimiser cross-entropy loss.
Corpus pretraining	Common Crawl (filtré), Wikipedia, livres, code (GitHub), articles scientifiques
Volume données	GPT-3: 300B tokens, GPT-4: ~13T tokens (estimé)
Compute pretraining	GPT-4: ~$100M de compute, ~25,000 A100 GPUs pendant 90-100 jours
Phase 2: SFT	Supervised Fine-Tuning sur conversations humain-AI annotées (~100K exemples)
Phase 3: RLHF	Reinforcement Learning from Human Feedback pour alignment
Reward model	Modèle entraîné sur comparaisons de réponses par annotateurs humains
PPO training	Proximal Policy Optimization pour ajuster le modèle selon reward model
Annotations RLHF	~1M comparaisons par annotateurs humains (Sama, Scale AI)
Constitutional AI	Anthropic approach : self-critique basée sur principes écrits
Red teaming	>50 experts externes testant les limites de sécurité avant lancement
5. CAPACITÉS ET BENCHMARKS	
MMLU	GPT-4: 86.4% (vs 70% GPT-3.5, vs ~90% humain expert)
HumanEval (code)	GPT-4: 67% pass@1 (vs 48% GPT-3.5)
Bar Exam	GPT-4: 90th percentile (vs 10th GPT-3.5)
SAT Math	GPT-4: 700/800 (93rd percentile)
GRE Writing	GPT-4: 4/6 (54th percentile)
Medical licensing (USMLE)	GPT-4: passe les 3 étapes avec marge significative
Coding assistance	Génération, debugging, refactoring, documentation, tests
Multilingual	~100 langues, performances corrélées à la représentation dans le corpus
Vision (GPT-4V)	Compréhension images : OCR, description, raisonnement visuel
Function calling	Génération structured output, appels API, agents
6. LIMITES STRUCTURELLES ET RISQUES	
Hallucinations	Génération d'informations fausses avec haute confiance. Taux : 5-20% selon domaine.
Mécanisme hallucination	Pas de grounding factuel : génère tokens probables, pas nécessairement vrais
Knowledge cutoff	Connaissances figées à la date d'entraînement (GPT-4: sept 2021 → updates)
Raisonnement math	Fragile sur calculs multi-étapes, arithmétique large nombres
Prompt injection	Manipulation du comportement via instructions cachées dans les inputs
Jailbreaking	Contournement des guardrails par prompts adversariaux
Sycophancy	Tendance à confirmer les croyances de l'utilisateur plutôt que corriger
Biais training data	Reproduit biais présents dans le corpus (genre, race, culture)
Misinformation	Capacité de générer de la désinformation convaincante à échelle
Deepfakes textuels	Imitation de style, génération de faux contenus attribués

Atrophie cognitive	Risque de délégation excessive réduisant les capacités humaines
Disruption emplois	McKinsey: 30% tâches automatisables, impact copywriters, support, code
7. USAGE ET MODÈLE ÉCONOMIQUE	
Pricing API (GPT-4)	Input: $30/1M tokens, Output: $60/1M tokens (list price 2023)
ChatGPT Plus	$20/mois, accès GPT-4, plugins, Code Interpreter
Enterprise	Pricing custom, data privacy renforcée, admin controls
Revenus OpenAI	~$1.6B ARR (2023), projection $5B (2024)
Coûts compute	Inference GPT-4 : ~$0.01-0.05 par requête (estimé)
Usages principaux	Coding (25%), Writing (20%), Research (15%), Learning (10%), Creative (10%)
Intégration Bing	10M+ DAU Bing Chat, parts de marché search ~3% → 4%
Copilot revenue	GitHub Copilot: $100/user/an, 1.5M users (2023)
8. GOUVERNANCE ET MESURES DE SÉCURITÉ	
Safety team	~30 chercheurs dédiés safety, + alignment team
Usage policies	Prohibited uses : armes, malware, spam, CSAM, deception, harassment
Content filtering	Classifier input/output pour détecter contenus policy-violating
Rate limiting	Limites par tier pour prévenir abus massifs
Monitoring	Analyse usage patterns pour détecter comportements suspects
Moderation API	API gratuite pour détecter contenus problématiques
Red teaming externe	Partenariats ARC, METR pour évaluation risques catastrophiques
Preparedness Framework	Cadre d'évaluation des risques avant déploiement (2023)
Evals open source	Publication benchmarks d'évaluation pour la communauté
Régulation	AI Act (EU) : GPT-4 = 'general purpose AI', obligations transparence
Voluntary commitments	White House (juillet 2023) : engagements sécurité, watermarking, red teaming

Section 3 — Matrices d'impact multidimensionnelles

Les matrices d'impact permettent une évaluation systématique des effets de chaque système sur un ensemble de dimensions standardisées. L'évaluation utilise une échelle à six niveaux : 1 (Négligeable), 2 (Faible), 3 (Modéré), 4 (Significatif), 5 (Majeur), C (Critique). Le niveau 'Critique' est réservé aux impacts de portée historique ou civilisationnelle.

Chaque évaluation est accompagnée d'une justification détaillée référençant les sources et les mécanismes causaux identifiés. Les matrices permettent la comparaison horizontale (un système sur plusieurs dimensions) et verticale (plusieurs systèmes sur une dimension).

3.1 Matrice d'impact — Google PageRank

GOOGLE PAGERANK — MATRICE D'IMPACT (1998-2012)		
DIMENSION	**SCORE**	**JUSTIFICATION DÉTAILLÉE**
Épistémologique	**5 - Majeur**	Redéfinition de la pertinence informationnelle. Citation comme proxy de qualité. Nouveau régime de vérité.
Fonctionnel	**5 - Majeur**	De la navigation manuelle à la recherche algorithmique. Transformation complète de l'accès à l'information.
Relationnel user-système	**4 - Significatif**	Interface minimaliste. Requête → résultats. Confiance implicite dans le ranking.
Économique direct	**5 - Majeur**	Création marché search ads ($150B+). Destruction annuaires. Restructuration médias.
Économique indirect	**5 - Majeur**	Industrie SEO ($65B). Dépendance éditeurs. Google Tax debate.
Social - pratiques info	**5 - Majeur**	'Googler' devient verbe. Transformation recherche quotidienne. Digital natives.
Social - inégalités	**4 - Significatif**	Fracture numérique. Avantage alphabétisés numériques. Biais linguistiques.
Démocratique - débat	**4 - Significatif**	Accès démocratisé à l'information. Mais concentration gatekeeping.
Démocratique - pouvoir	**4 - Significatif**	Nouveau pouvoir privé sur l'accès au savoir. Pas de contrôle démocratique.
Environnemental	**3 - Modéré**	Datacenters massifs. Empreinte carbone croissante. Mais efficacité accès info.

3.2 Matrice d'impact — Facebook News Feed

FACEBOOK NEWS FEED — MATRICE D'IMPACT (2006-2022)		
DIMENSION	**SCORE**	**JUSTIFICATION DÉTAILLÉE**
Épistémologique	**5 - Majeur**	La vérité = ce qui engage. Glissement vers post-vérité. Viralité ≠ véracité.
Fonctionnel	**5 - Majeur**	De l'information choisie à l'information poussée. Passivité consommation.
Relationnel user-système	**5 - Majeur**	Boucles de rétroaction. Addiction par design. Notifications. Variable rewards.

Économique direct	5 - Majeur	Marché social ads $150B+. ARPU $40 US. Domination attention economy.
Économique indirect	5 - Majeur	Destruction médias traditionnels. Précarisation journalisme. Creator economy.
Social - pratiques info	CRITIQUE	CRITIQUE. Restructuration espace public. Bulles de filtre. Polarisation mesurée.
Social - inégalités	4 - Significatif	Micro-targeting publicitaire. Discrimination algorithmique documentée (housing, jobs).
Social - santé mentale	5 - Majeur	Corrélations anxiété/dépression ados. Études internes Facebook Files.
Démocratique - débat	CRITIQUE	CRITIQUE. Cambridge Analytica. Désinformation électorale. Érosion confiance.
Démocratique - pouvoir	5 - Majeur	Pouvoir arbitraire sur discours. Modération opaque. XCheck VIP system.
Droits humains	5 - Majeur	Myanmar, Éthiopie : rôle dans violence ethnique documenté. Aveux tardifs.

3.3 Matrice d'impact — GPT-4

GPT-4 / CHATGPT — MATRICE D'IMPACT (2020-2024)		
DIMENSION	**SCORE**	**JUSTIFICATION DÉTAILLÉE**
Épistémologique	CRITIQUE	CRITIQUE. Génération vs révélation. Hallucinations normalisées. Vérité = cohérence.
Fonctionnel	CRITIQUE	CRITIQUE. De chercher à déléguer. Atrophie cognitive potentielle. Dépendance.
Relationnel user-système	5 - Majeur	Anthropomorphisation. Relation conversationnelle. Confiance excessive.
Économique direct	5 - Majeur	Marché GenAI $40B+ (2024). Disruption SaaS. Concentration OpenAI/Google/Anthropic.
Économique - emploi	CRITIQUE	CRITIQUE. McKinsey: 30% tâches automatisables. Copywriters, support, coding, legal.
Social - éducation	5 - Majeur	Plagiat généralisé. Refonte évaluation. Disruption apprentissage écriture.
Social - création	5 - Majeur	Génération massive contenu. Dévaluation créativité humaine. Copyright chaos.
Démocratique - info	5 - Majeur	Désinformation à échelle industrielle. Deepfakes textuels. Manipulation facilitée.
Démocratique - pouvoir	5 - Majeur	Concentration : 3 acteurs dominent. Coûts training = barrière. Pas de contrôle public.
Sécurité - dual use	5 - Majeur	Phishing, malware, manipulation sociale facilités. Biorisques potentiels.
Existentiel (débattu)	4 - Significatif	Risques alignement long-terme. Débat experts. AI Safety priorité OpenAI.

Section 4 — Chronologie annotée (1993-2025)

La chronologie rassemble 60+ événements clés de l'histoire des algorithmes de l'information. Chaque événement est classé selon son type (Architecturale, Paramétrique, Opérationnelle, Contextuelle) et sa catégorie (Search, Social, AI, Policy, Business, Academia). Les descriptions synthétisent l'impact et la signification de chaque événement dans la trajectoire d'ensemble.

4.1 Années 1990-2000 : Fondations

DATE	ÉVÉNEMENT	TYPE	CAT.	IMPACT
1993	**Excite, premier moteur par ML**	Architecturale	Search	Statistiques sur fréquence mots. Précurseur IA dans search.
1994	**Yahoo! Directory**	Opérationnelle	Search	Curation humaine. Modèle éditorial vs algorithmique.
1995	**AltaVista**	Architecturale	Search	Premier index full-text massif. Référence pré-Google.
1996	**BackRub (proto-PageRank)**	Architecturale	Search	Page & Brin à Stanford. Analyse des backlinks.
1998	**Google Search / PageRank**	Architecturale	Search	RUPTURE MAJEURE. Citation-based ranking. Refonde le search.
1998	**Publication Brin & Page WWW7**	Contextuelle	Academia	Paper fondateur. Diffusion académique du concept.

4.2 Années 2000-2010 : Expansion

DATE	ÉVÉNEMENT	TYPE	CAT.	IMPACT
2000	**Google AdWords**	Opérationnelle	Business	Publicité search. Modèle économique dominant du web.
2000	**Bulle internet / crash**	Contextuelle	Économie	Consolidation du marché. Google survivant majeur.
2002	**Google News**	Opérationnelle	Search	Agrégation algorithmique actualités. Début tension avec médias.
2003	**Florida Update**	Paramétrique	Search	Première major update. Pénalise keyword stuffing. Séisme SEO.
2004	**Facebook lancé**	Contextuelle	Social	Harvard → monde. Nouveau paradigme social.
2004	**Gmail (2GB gratuit)**	Contextuelle	Tech	Scanning emails pour ads. Data collection massive.
2004	**Google IPO**	Contextuelle	Business	$23B valorisation. Légitimation du modèle.
2005	**Google Maps / Earth**	Opérationnelle	Search	Cartographie algorithmique. 'La carte devient le territoire.'
2006	**Twitter lancé**	Contextuelle	Social	Microblogging. Temps réel. Nouveau format informationnel.

2006	**Facebook News Feed**	Architecturale	Social	RUPTURE. Premier flux algorithmique de masse. EdgeRank.
2006	**Réaction anti-News Feed**	Contextuelle	Social	Protest users 'Stalking!' Première techlash sur feed.
2007	**iPhone**	Contextuelle	Tech	Mobile-first. Ubiquité. GPS + toujours connecté.
2007	**Universal Search**	Opérationnelle	Search	Blending résultats (images, vidéos, news, maps) dans SERP.
2008	**App Store iOS**	Contextuelle	Tech	Écosystème apps. Nouveau canal distribution.
2008	**Android**	Contextuelle	Tech	Démocratisation smartphone. Google mobile.
2009	**Personalized Search Google**	Opérationnelle	Search	Fin résultats universels. Historique utilisateur. Bulles.
2009	**Facebook Like button**	Opérationnelle	Social	Signal d'engagement universel. Tracking web entier.

4.3 Années 2010-2020 : Personnalisation et crises

DATE	ÉVÉNEMENT	TYPE	CAT.	IMPACT
2010	**Caffeine (Google)**	Architecturale	Search	Nouvelle infra indexation. 50% plus frais.
2010	**Instagram lancé**	Contextuelle	Social	Photo-first social. Visual culture.
2011	**Google Panda**	Paramétrique	Search	Qualité contenu. Pénalise thin content, content farms. 12% queries.
2011	**Filter Bubble (Pariser)**	Contextuelle	Academia	Livre. Conceptualisation des bulles de filtre. Débat public.
2011	**Facebook ML ranking**	Architecturale	Social	EdgeRank → Machine Learning. Opacification.
2012	**Google Penguin**	Paramétrique	Search	Qualité liens. Pénalise link schemes. 3% queries.
2012	**Knowledge Graph**	Opérationnelle	Search	'Strings to things.' Réponses directes. Début zero-click.
2012	**AlexNet / ImageNet**	Architecturale	AI	RUPTURE DEEP LEARNING. CNN révolution vision. Début ère moderne AI.
2012	**Facebook 1 milliard users**	Contextuelle	Social	Échelle sans précédent. Infrastructure sociale mondiale.
2012	**Instagram acquis $1B**	Contextuelle	Business	Consolidation Meta. Antitrust future.
2013	**Snowden / PRISM**	Contextuelle	Policy	Révélation surveillance. Techlash v1. Chiffrement.
2013	**Word2Vec (Google)**	Architecturale	AI	Embeddings mots. Fondation NLP moderne.
2014	**Facebook Emotional Contagion**	Contextuelle	Research	Expérience manipulation mood 700K users sans consentement. Scandale.
2014	**GAN (Goodfellow)**	Architecturale	AI	Generative Adversarial Networks. Génération réaliste.

2014	**Amazon Alexa**	Opérationnelle	AI	Voice assistant. Nouvelle interface NL.
2015	**RankBrain (Google)**	Architecturale	Search	ML dans ranking core. Opacité. 3ème signal importance.
2015	**Deep learning mainstream**	Contextuelle	AI	TensorFlow open source. Démocratisation.
2016	**Élection US / micro-targeting**	Contextuelle	Policy	Rôle réseaux sociaux débattu. Désinformation russe.
2016	**Google Duplex demo**	Opérationnelle	AI	AI passe pour humain au téléphone. Débat éthique.
2016	**AlphaGo bat Lee Sedol**	Contextuelle	AI	Moment symbolique. Deep RL surpasse humains sur Go.
2017	**Transformer paper**	Architecturale	AI	RUPTURE FONDATRICE LLM. 'Attention is all you need.' Architecture dominante.
2018	**Cambridge Analytica**	Contextuelle	Policy	SCANDALE MAJEUR. 87M users. Congressional hearings. RGPD accéléré.
2018	**RGPD entré en vigueur**	Contextuelle	Policy	Première régulation data majeure. Modèle mondial.
2018	**Facebook MSI update**	Paramétrique	Social	Meaningful Social Interactions. Réponse aux critiques.
2018	**BERT (Google)**	Architecturale	AI	Bidirectional transformers. Boost NLU search.
2018	**GPT-1 (OpenAI)**	Architecturale	AI	Premier GPT. 117M params. Proof of concept.
2019	**GPT-2 (OpenAI)**	Architecturale	AI	1.5B params. Release staged pour 'risques.' Médiatisation.
2019	**FTC amende Facebook $5B**	Contextuelle	Policy	Plus grosse amende FTC histoire. Consent decree.

4.4 Années 2020-2025 : Rupture générative

DATE	ÉVÉNEMENT	TYPE	CAT.	IMPACT
2020	**GPT-3 (OpenAI)**	Architecturale	AI	RUPTURE. 175B params. Capacités émergentes. Few-shot learning.
2020	**COVID + désinformation**	Contextuelle	Policy	Infodémie. Pression modération. Fact-checking massif.
2020	**Facebook Oversight Board**	Opérationnelle	Social	Gouvernance externe. Quasi-judicial. Expérimentation.
2021	**DALL-E (OpenAI)**	Architecturale	AI	Génération images depuis texte. Multimodal génératif.
2021	**Facebook Files (WSJ)**	Contextuelle	Policy	Fuite documents internes. Savait impacts négatifs. Techlash v2.
2021	**Facebook → Meta**	Contextuelle	Business	Rebranding. Pivot metaverse. Fuite en avant.
2021	**Chinchilla scaling laws**	Architecturale	AI	DeepMind. Optimal params/data ratio. Influence training.
2022	**ChatGPT lancé**	Opérationnelle	AI	RUPTURE MAJEURE. LLM grand public. 100M users 2 mois. Choc sociétal.
2022	**Stable Diffusion**	Architecturale	AI	Image génération open source. Démocratisation.
2022	**Twitter → X (Musk)**	Contextuelle	Social	Acquisition. Chaos. Algorithme open source (partiel).
2023	**GPT-4 lancé**	Architecturale	AI	Multimodal. Performance near-human. Benchmarks records.
2023	**Bing Chat / New Bing**	Opérationnelle	Search	Microsoft + OpenAI. LLM dans search. Fin des liens bleus.
2023	**Google SGE / Bard**	Opérationnelle	Search	Réponse Google à ChatGPT. Search Generative Experience.
2023	**Lettre pause AI (Future of Life)**	Contextuelle	Policy	1,000+ signataires. Débat risques existentiels.
2023	**EU AI Act voté**	Contextuelle	Policy	Première régulation AI contraignante. High-risk categories.
2023	**Executive Order AI (US)**	Contextuelle	Policy	Biden. Standards sécurité. Reporting obligations.
2023	**Claude 2 (Anthropic)**	Architecturale	AI	Concurrent majeur. Constitutional AI. Focus safety.
2023	**Llama 2 (Meta)**	Architecturale	AI	Open weights. Démocratisation. Écosystème open source.
2024	**GPT-4 Turbo 128K**	Paramétrique	AI	Context window massif. Prix réduits. Mainstream.
2024	**Sora (OpenAI)**	Architecturale	AI	Génération vidéo. Prochain frontier.
2024	**Google Gemini**	Architecturale	AI	Réponse Google. Multimodal native. Compétition.
2024	**Claude 3 (Anthropic)**	Architecturale	AI	Opus approche GPT-4. Pluralité marché LLM.

2024	**AI Act entre en vigueur**	Contextuelle	Policy	Application progressive. Obligations concrètes.
2025	**Agents autonomes mainstream**	Opérationnelle	AI	De la génération à l'action. Computer use. Délégation complète.

Section 5 — Grilles d'audit algorithmique

Les grilles d'audit constituent des outils opérationnels pour l'évaluation de conformité des systèmes algorithmiques. Elles sont structurées en six domaines : Transparence (A), Équité (B), Responsabilité (C), Sécurité (D), Données (E), Impact (F). Chaque critère peut être évalué comme Conforme, Partiel, ou Non-conforme, avec documentation des observations et preuves.

Ces grilles sont alignées avec les principales régulations et standards : AI Act (UE), RGPD, ISO/IEC 42001 (AI Management), NIST AI RMF, IEEE 7000. Elles peuvent être adaptées aux contextes sectoriels spécifiques.

5.1 Grille d'audit complète

GRILLE D'AUDIT ALGORITHMIQUE — CHECKLIST EXHAUSTIVE			
CRITÈRE	**CONFORME**	**PARTIEL**	**OBSERVATIONS / PREUVES**
A. TRANSPARENCE			
A1. Documentation publique de l'existence du système	☐	☐	
A2. Description intelligible du fonctionnement général	☐	☐	
A3. Publication des critères de décision principaux	☐	☐	
A4. Information des utilisateurs sur l'usage d'algorithmes	☐	☐	
A5. Explication individuelle des décisions sur demande	☐	☐	
A6. Publication des mises à jour significatives	☐	☐	
A7. Accès chercheurs aux données d'analyse	☐	☐	
B. ÉQUITÉ ET NON-DISCRIMINATION			
B1. Analyse des biais sur attributs protégés	☐	☐	
B2. Tests de disparate impact	☐	☐	
B3. Mécanismes de correction des biais identifiés	☐	☐	
B4. Diversité des données d'entraînement	☐	☐	

B5. Équité des outcomes mesurée	☐	☐	
B6. Accessibilité pour utilisateurs handicapés	☐	☐	
C. RESPONSABILITÉ ET GOUVERNANCE			
C1. Responsable identifié pour le système	☐	☐	
C2. Chaîne de responsabilité documentée	☐	☐	
C3. Procédure de recours pour les affectés	☐	☐	
C4. Mécanisme d'escalade des incidents	☐	☐	
C5. Audit interne régulier	☐	☐	
C6. Audit externe indépendant	☐	☐	
C7. Comité éthique consultatif	☐	☐	
D. SÉCURITÉ ET ROBUSTESSE			
D1. Tests de robustesse adversariale	☐	☐	
D2. Protection contre les attaques par injection	☐	☐	
D3. Mécanismes de fallback en cas d'échec	☐	☐	
D4. Monitoring des comportements anormaux	☐	☐	
D5. Procédure de désactivation d'urgence	☐	☐	
D6. Red teaming régulier	☐	☐	
E. DONNÉES ET VIE PRIVÉE			
E1. Minimisation des données collectées	☐	☐	
E2. Consentement explicite pour données sensibles	☐	☐	
E3. Data retention policy claire	☐	☐	
E4. Droit d'accès aux données implémenté	☐	☐	

E5. Droit à l'effacement implémenté	☐	☐	
E6. Portabilité des données	☐	☐	
E7. Anonymisation/pseudonymisation	☐	☐	
F. ÉVALUATION D'IMPACT			
F1. Étude d'impact réalisée avant déploiement	☐	☐	
F2. Consultation parties prenantes	☐	☐	
F3. Mesure des impacts post-déploiement	☐	☐	
F4. Mécanisme de feedback utilisateurs	☐	☐	
F5. Révision périodique de l'impact	☐	☐	

Section 6 — Synthèse comparative des trois paradigmes

Cette section présente une analyse transversale des trois systèmes étudiés, mettant en évidence les continuités et les ruptures entre les paradigmes citationnel, comportemental et génératif.

6.1 Tableau comparatif des dimensions structurantes

DIMENSION	PAGERANK (Citationnel)	NEWS FEED (Comportemental)	GPT-4 (Génératif)
Épistémologie	Confiance collective via citations	Prédiction comportementale	Cohérence discursive
Fonction centrale	Révéler l'existant	Prédire les préférences	Générer le nouveau
Relation utilisateur	Consultation active	Consommation passive	Délégation cognitive
Source de vérité	Le réseau des pairs	Le comportement passé	Le modèle lui-même
Temporalité	Accumulation (liens)	Instantané (engagement)	Génération à la demande
Transparence	Principe public, détails secrets	Opaque (ML black box)	Très opaque (MoE, RLHF)
Risque principal	Concentration (Matthieu)	Polarisation (bulles)	Hallucination (déréalisation)
Vulnérabilité	Link manipulation	Engagement gaming	Prompt injection
Modèle économique	Ads search ($150B+)	Ads social ($150B+)	SaaS + API ($5B+)
Concentration	Google 90%+	Meta 70%+ (social)	OpenAI/Google/Anthropic
Régulation applicable	Antitrust, concurrence	DSA, protection données	AI Act, dual-use
Accountability	Limitée (secret industriel)	Faible (opacity)	Émergente (standards)
Disruption emploi	Faible directe	Modérée (médias)	Forte (30% tâches)

6.2 Évolution des risques entre paradigmes

Le graphique ci-dessous illustre l'évolution des cinq axes de risque principaux à travers les trois paradigmes. On observe une tendance générale à l'augmentation de l'opacité et de la dépendance, tandis que les risques de manipulation et de biais présentent des profils différenciés selon les paradigmes.

AXE	PR (1)	PR (2)	PR (3)	FB (1)	FB (2)	GPT
Opacité	2 - Faible	3 - Modéré	3 - Modéré	4 - Significatif	5 - Majeur	5 - Majeur
Biais	2 - Faible	3 - Modéré	3 - Modéré	4 - Significatif	5 - Majeur	4 - Significatif
Manipulation	3 - Modéré	4 - Significatif	3 - Modéré	4 - Significatif	5 - Majeur	4 - Significatif
Dépendance	3 - Modéré	4 - Significatif	4 - Significatif	4 - Significatif	5 - Majeur	5 - Majeur

Irréversibilité	**2 - Faible**	**3 - Modéré**	**3 - Modéré**	**3 - Modéré**	**4 - Significatif**	**4 - Significatif**

Légende : PR = PageRank phases (1: 1998-2005, 2: 2005-2010, 3: 2010-2012) ; FB = Facebook (1: EdgeRank, 2: ML) ; GPT = 2022-2024

Section 7 — Analyses sectorielles

Cette section examine les impacts spécifiques des mutations algorithmiques dans six secteurs clés : Médias et journalisme, Éducation, Santé, Emploi, Commerce, Démocratie. Chaque analyse sectorielle identifie les mécanismes d'impact, les acteurs affectés, et les enjeux de gouvernance propres au secteur.

7.1 Médias et journalisme

IMPACT SUR LES MÉDIAS ET LE JOURNALISME	
Mécanisme clé PageRank	Concentration du trafic sur quelques sources. 'Portail' unique vers l'information.
Mécanisme clé Facebook	Dépendance referral (30-50% trafic). Pivot to video désastreux (2016-2018). Effondrement reach.
Mécanisme clé LLM	Génération de résumés sans visite source. Zero-click search. Désintermédiation totale.
Emplois affectés	Journalistes -30% (US 2008-2020), photographes, fact-checkers, rédacteurs SEO.
Modèle économique	Effondrement pub print. Transition abo difficile. Concentration groupes (GAFAM acheteurs).
Qualité information	Course au clic. Simplification. Polarisation. Désinformation virale.
Pluralisme	Concentration éditoriale. Déserts informationnels locaux. Uniformisation traitement.
Régulation spécifique	Droits voisins (EU), News Media Bargaining Code (AU), négociations Google News Showcase.
Enjeu LLM	Scraping copyright. Compensation créateurs. Traçabilité sources. Hallucination factuelle.

7.2 Éducation

IMPACT SUR L'ÉDUCATION	
Accès connaissances	Démocratisation (Wikipedia, Khan Academy, MOOCs). Mais filter bubbles sur sujets controversés.
Plagiat pré-LLM	Copy-paste détectable. Turnitin. Banques de travaux. Paraphrase manuelle.
Plagiat post-LLM	Génération originale indétectable. Crise de l'évaluation écrite. Refonte pédagogique urgente.
Apprentissage écriture	Risque d'atrophie. Délégation précoce. Perte compétence rédactionnelle.
Inégalités	Premium tools payants (GPT-4 $20/mois). Fracture numérique. Avantage anglophones.
Tutoring personnalisé	Opportunité : aide 24/7, adaptation rythme. Risque : dépendance, hallucinations.
Réponses institutions	Interdictions (NYC), intégration curriculaire, évaluation orale, projects hands-on.
Compétences futures	Prompt engineering. Vérification critique. Collaboration humain-IA. Meta-cognition.

7.3 Emploi

IMPACT SUR L'EMPLOI	
Estimation automatisation	McKinsey: 30% tâches automatisables par GenAI. Goldman: 300M emplois affectés mondial.
Secteurs très exposés	Support client, rédaction, traduction, coding junior, legal research, data entry.
Secteurs moins exposés	Soins (empathie), artisanat, leadership, négociation, créativité haute.
Augmentation vs remplacement	Débat : productivité +40% (études GitHub Copilot) vs réduction effectifs.
Gig economy	Algorithmes de dispatch (Uber, DoorDash). Précarisation. Surveillance performance.
Recrutement algorithmique	Screening CV, matching. Biais documentés (Amazon 2018). Opacité critères.
Formation/reconversion	Urgence upskilling. Mais vitesse changement > capacité adaptation.
Régulation travail	Directive EU plateformes. Droit à l'explication décisions algorithmiques.

Section 8 — Recommandations opérationnelles

Cette section formule des recommandations concrètes destinées aux trois catégories d'acteurs concernés par les mutations algorithmiques : organisations déployant des systèmes, régulateurs et pouvoirs publics, utilisateurs et société civile.

8.1 Recommandations aux organisations

- Mettre en place une gouvernance algorithmique formalisée avec responsable identifié et comité éthique.
- Réaliser des études d'impact algorithmique (AIA) avant tout déploiement significatif.
- Documenter systématiquement les systèmes via des 'model cards' ou fiches algorithmiques standardisées.
- Tester les biais sur attributs protégés et mettre en place des mécanismes de correction.
- Établir des procédures de recours et d'explication pour les personnes affectées.
- Former les équipes à la littératie algorithmique et aux enjeux éthiques.
- Prévoir des mécanismes de désactivation et de fallback en cas de dysfonctionnement.
- Participer aux initiatives sectorielles de standardisation et de bonnes pratiques.

8.2 Recommandations aux régulateurs

- Développer l'expertise technique interne pour réduire l'asymétrie avec les régulés.
- Adopter une approche principle-based permettant l'adaptation aux évolutions technologiques.
- Exiger la transparence sur l'existence et les effets des systèmes algorithmiques.
- Mettre en place des mécanismes d'audit indépendant avec accès aux données nécessaires.
- Créer des sandboxes réglementaires pour expérimenter les nouvelles approches.
- Renforcer la coopération internationale face à des systèmes transfrontaliers.
- Soutenir la recherche indépendante sur les impacts algorithmiques.
- Développer des standards de certification et de labellisation.

8.3 Recommandations aux utilisateurs et à la société civile

- Développer la littératie algorithmique : comprendre les logiques des systèmes utilisés.
- Diversifier les sources d'information pour contrer les bulles de filtre.
- Utiliser les outils de contrôle disponibles (paramètres vie privée, préférences).
- Exercer les droits existants (accès, rectification, explication, portabilité).
- Soutenir les initiatives de transparence et d'audit citoyen.
- Participer aux consultations publiques sur la régulation algorithmique.
- Maintenir une capacité de jugement critique face aux outputs générés.
- Préserver des espaces de friction cognitive contre la délégation totale.

Conclusion

Cette cartographie exhaustive a documenté les trois mutations majeures qui ont transformé les systèmes algorithmiques de l'information au cours des trois dernières décennies. Du paradigme citationnel du PageRank au paradigme comportemental des réseaux sociaux, puis au paradigme génératif des grands modèles de langage, chaque mutation a reconfiguré notre rapport à l'information, à la connaissance et à la réalité.

Les outils rassemblés dans ce document — fiches techniques, matrices d'impact, chronologies, grilles d'audit, analyses sectorielles — constituent un arsenal méthodologique pour comprendre, évaluer et gouverner ces systèmes. Ils sont conçus pour être utilisés par les praticiens confrontés à la complexité des mutations algorithmiques : auditeurs, régulateurs, chercheurs, décideurs, citoyens engagés.

L'enjeu fondamental qui traverse l'ensemble de cette cartographie est celui de la souveraineté cognitive : la capacité des individus et des collectivités à conserver un degré d'autonomie dans un monde où l'accès à l'information est massivement intermédié par des systèmes automatisés. Préserver cette souveraineté suppose une vigilance constante, une compréhension des logiques algorithmiques, et une participation active à leur gouvernance.

Les mutations algorithmiques ne sont pas des phénomènes naturels face auxquels nous serions impuissants. Ce sont des constructions sociotechniques, produites par des choix humains, répondant à des objectifs identifiables, susceptibles d'être interrogées et transformées. La cartographie proposée ici vise précisément à rendre ces systèmes intelligibles, à les soustraire à l'opacité qui les protège, à ouvrir l'espace d'une délibération collective sur leur devenir.

— FIN DU DOCUMENT —

OBSERVIA — Cartographie Exhaustive v1.0 — Janvier 2026