

Programme de Recherche

Gestion, Gouvernance et Ingénierie des Systèmes d'Intelligence Artificielle

Vision générale et positionnement

Après vingt ans de travaux à l'intersection des systèmes d'information, des données et de la gestion, ma recherche s'inscrit dans la gouvernance et la sécurité des systèmes d'IA. La question centrale n'est plus « que peut-on faire avec l'IA ? », mais « comment la gouverner, l'exploiter et la sécuriser responsablement ? »

Ce programme s'appuie sur trois blocs d'expérience : (1) ancrage académique en gouvernance d'IA avec publications dans *Ethics and Information Technology* et *Government Information Quarterly*; (2) pratique terrain de projets d'IA dans télécom, énergie, santé et secteur public; (3) cartographie de 250 000+ brevets en IA/ML (2010–2025).

Trois axes de recherche complémentaires

Axe 1 – Gouvernance et gestion des risques de l'IA

Mes publications antérieures (« AI Regulation », 2021; « AI Governance in Public Organizations », 2025) ont établi des cadres de régulation et documenté les pratiques dans 28 organisations publiques sur cinq continents. Cet axe prolonge ce travail au niveau organisationnel et des portefeuilles d'IA, avec trois volets complémentaires :

- **Structuration de la gouvernance** : Quels rôles (Chief AI Officer, comités d'éthique, équipes de conformité), quels comités (exécutif, risques, innovation), quels processus de décision (approvals, escalade) permettent une gouvernance effective? Comment adapter ces structures selon le secteur, la maturité numérique et la taille de l'organisation?
- **Cartographie et gestion des risques** : Au-delà des risques classiques (biais, conformité réglementaire), nous documenterons les risques émergents spécifiques aux LLM et agents autonomes : prompt injection, hallucinations incontrôlées, dépendance aux fournisseurs de modèles, perte de contrôle sur des agents agentiques, boucles affectives incontrôlées, drift de modèles, contamination de données d'entraînement. Comment créer des taxonomies de risques exploitables et des cartes de contrôles associés?
- **Contrôle réversible et adaptabilité** : Comment les organisations peuvent-elles implémenter un « contrôle réversible » permettant d'ajuster dynamiquement leurs systèmes d'IA selon l'évolution des risques et des apprentissages, tout en maintenant des capacités d'apprentissage organisationnel et de pivotage rapide?

Méthodologie mixte : études de cas approfondies (8–10 organisations), entretiens semi-structurés, analyse de documents (politiques, registres de risques), enquêtes quantitatives auprès de responsables IA.

Livrables attendus : (1) typologies de structures de gouvernance par secteur et maturité; (2) taxonomie de risques de l'IA et portefeuille de contrôles; (3) guide pratique de mise en place; (4) 2–3 cas d'enseignement pour le microprogramme de 3e cycle en gestion stratégique de l'IA.

Axe 2 – Ingénierie, opérations et sécurité des systèmes d'IA (AIOps/MLOps)

Ce volet se situe au cœur du cycle de vie des systèmes d'IA : collecte et gouvernance des données, construction et évaluation de modèles, déploiement, exploitation en production, monitoring, maintenance, sécurité continue. C'est le terrain où les principes de gouvernance se matérialisent réellement. Quatre domaines de recherche prioritaires :

- **Pipelines AIOps/MLOps gouvernables** : Comment concevoir des pipelines où traçabilité, contrôles d'accès, revues de risques, audits, mécanismes d'arrêt d'urgence (« kill switches ») et points de décision humaine font partie intégrante de l'architecture et des outils, plutôt que de rester dans des documents? Comment associer gouvernance et performance/rapidité?
- **Patrons d'architecture pour systèmes à base de LLM** : Quelles architectures de référence pour RAG (Retrieval-Augmented Generation), chaînes d'outils, agents multi-étapes permettent de garder la main sur ce que le système peut faire, sur quelles données, avec quelles garanties de sécurité et performance? Comment intégrer des « agents firewall » encadrant les actions des LLM?
- **Sécurité des LLM et atténuation des risques** : (a) Conception d'un « pare-feu LLM » multi-phases : analyse de brevets pour identifier les innovations, conception de garde-fous au niveau de l'input (prompt injection), du modèle (finetune sécurisé), et de l'output (validation et filtering). (b) Évaluation et atténuation des hallucinations critiques dans les systèmes RAG : quand sont-elles dangereuses, comment les détecter et les réduire? (c) Sécurité des agents : contrôle des permissions, audit des actions, limite de boucles de raisonnement.

- **Boucles affectives en interaction humain-IA** : Comment émergent les boucles affectives dans l'interaction prolongée avec des systèmes conversationnels? Quelles implications pour la conception des systèmes, l'analyse de sentiment, et la responsabilité organisationnelle?

Méthodologie : design science et recherche orientée artefact. Développement et évaluation de prototypes de pipelines MLOps intégrant gouvernance, systèmes RAG avec atténuation d'hallucinations, agents firewall, tableaux de bord opérationnels. Expériences avec organisations partenaires (secteur public, télécom, énergie, santé).

Livrables attendus : (1) architectures de référence et patrons de design; (2) prototypes de pare-feu LLM et systèmes RAG sécurisés; (3) cadres d'évaluation de sécurité et performance; (4) publications : *Communications of the ACM, Academy of Management Review, MISQ*; (5) cas d'enseignement avancés et ressources open-source.

Axe 3 – Performance, impact et durabilité de l'IA

Cet axe répond à une question centrale que j'entends régulièrement des gestionnaires : l'IA crée-t-elle réellement de la valeur, pour qui et à quel coût? Le passage de la preuve de concept à la production révèle des écarts significatifs. Trois sous-domaines :

- **Mesure multi-dimensionnelle de la performance** : Au-delà des métriques techniques (précision, rappel, F1), comment mesurer la valeur réelle générée (ROI, impact sur les processus, efficacité opérationnelle), la qualité de service, l'équité des décisions, et l'impact sur le travail humain (déplacement d'emplois vs augmentation des capacités)?
- **Facteurs de succès et passage à l'échelle** : Pourquoi certains projets passent-ils à l'échelle organisationnelle tandis que d'autres restent bloqués au stade pilote? Analyse des facteurs : alignement stratégique, maturité des données, structures de gouvernance, capacités AIOps/MLOps, acceptation des utilisateurs, leadership organisationnel.
- **Durabilité environnementale, économique et sociale** : Comment intégrer des critères de durabilité (coûts énergétiques de l'entraînement et inférence, équité d'accès, impact social) dans la priorisation et l'évaluation des portefeuilles d'IA? Comment construire des systèmes d'IA résilients et adaptables?

Méthodologie : études de cas longitudinales (avant/après, quasi-expériences), analyses de séries temporelles, tableaux de bord co-conçus avec partenaires.

Livrables attendus : (1) frameworks d'évaluation d'impact multi-dimensionnels; (2) guide de priorisation de portefeuilles d'IA; (3) 4–5 cas d'études d'impact publiés; (4) ressources pédagogiques pour les formations de l'École de gestion auprès des gestionnaires et professionnels.

Programmation sommaire

- **Année 1** : Consolidation des partenariats, demandes de financement (CRSH, FRQSC, Mitacs), études pilotes.
- **Années 2–3** : Études de cas approfondies (8–10 organisations), prototypes technologiques, publications initiales, 2 cas d'enseignement.
- **Années 4–5** : Synthèse comparative, frameworks et taxonomies finalisés, dissémination praticiens, partenariats long terme, articles synthèse.

Intégration, formation et réseautage

Ce programme s'aligne avec quatre thèmes prioritaires de l'École : (1) transformation numérique et innovation, (2) analytique et données, (3) cybersécurité et résilience, (4) durabilité et impact social. Le SIMQG, CROI et Centre Laurent Beaudoin sont des partenaires naturels.

Formation HQP : 4–5 sujets maîtrise/docteurat, stages appliqués (6–8 mois), 4–5 cas d'enseignement pour microprogramme 3e cycle. Objectif à 5 ans : pôle de référence régional et international.

Financement : CRSH (Savoir, Partenariat), FRQSC, Mitacs, fondations privées, partenariats directs (secteur public, télécom, énergie, santé).

Dissémination : Revues de premier plan (*Communications of the ACM, Academy of Management Review, MISQ*), conférences majeures (ICIS, HICSS, AOM), webinaires praticiens, guides open-source, standards (ISO, cadres publics).