

I. Diagnostic Structuré de la Crise

État de la situation :

Trois modèles en production présentent des divergences métriques majeures. L'analyse révèle une dégradation profonde de la qualité et de l'éthique, masquée par une amélioration artificielle des performances de 93.1% (x14.5 plus rapide) enregistrée le 12 janvier 2024.

Source du problème :

La crise résulte d'une convergence de trois facteurs :

1. Technique : Migration vers un nouveau dataset mal maîtrisée et fuite de données temporelle (Ticket DATA-889).
2. Malveillance : Sabotage caractérisé par Alex (arrondi agressif dans `train_utils.py`, intrusions SSH nocturnes entre 22h et 4h).
3. Gouvernance : Pression business (BIZ-445) ayant conduit à des revues de code expéditives (3 PRs approuvées en < 1h) et une négligence des tests de fairness.

II. Hypothèses et Méthode d'Investigation

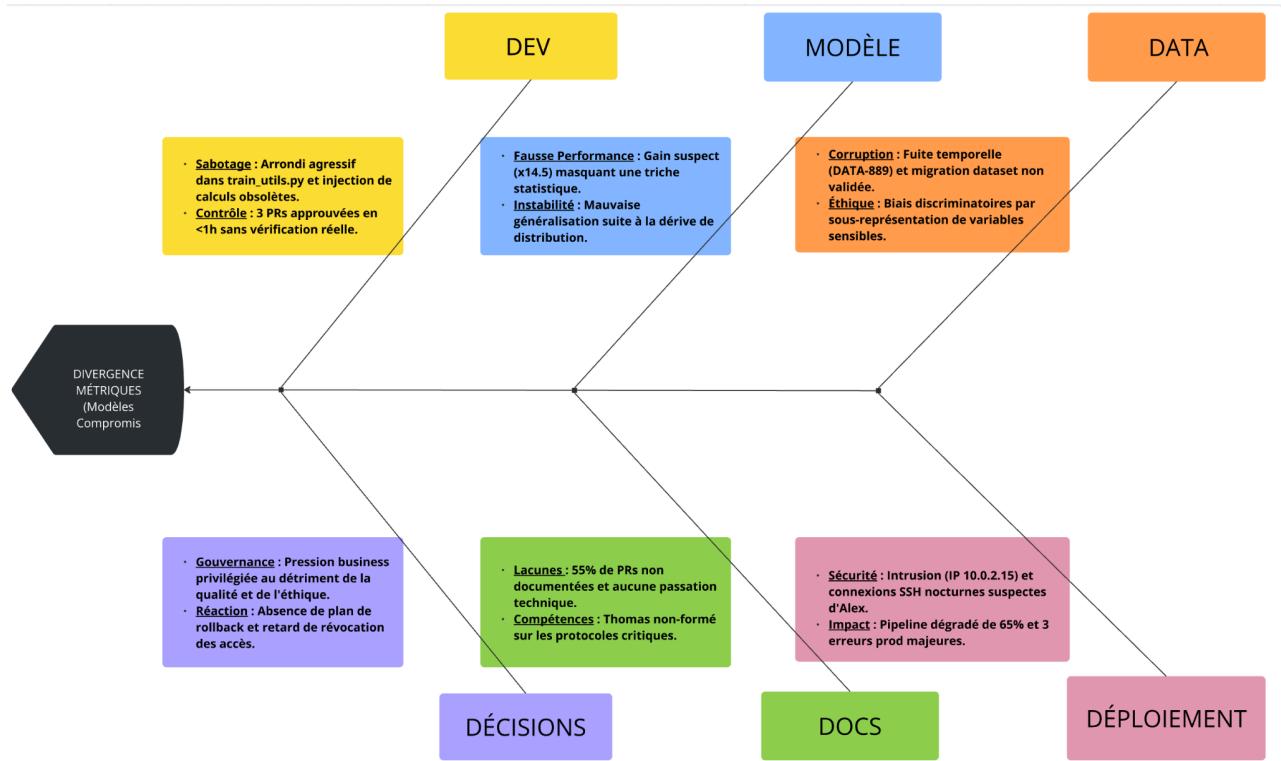
Méthode : Diagramme d'Ishikawa (6 axes)

Classification des Hypothèses

Identifiant	Hypothèse	Probabilité	Impact	Statut
H2	Dérive des features sensibles (âge, genre)	Élevée	Critique	Confirmé : Violation fairness (Cas Marie), importance feature +353%.
H1	Migration dataset non revalidée	Élevée	Élevé	Confirmé : Ticket DATA-889, dérive conceptuelle détectée.

H4	Faiblesse des tests et environnements	Élevée	Élevé	Confirmé : Tickets DEV-778 et INFRA-223.
H3	Métriques ajustées masquant la réalité	Moyenne	Élevé	Confirmé : $\$β=1.08\$, gain suspect x14.5$ (triche statistique).
H5	Sécurité / accès prod	Faible	Très élevé	Confirmé : Logs SSH suspects d'Alex (22h-4h), IP 10.0.2.15.

Méthode : Diagramme d'Ishikawa (6 axes)



III. Éléments Confirmés vs Écartés (Preuves)

- Confirmé : Sabotage du pipeline.
 - Preuve : L'arrondi trop agressif dans `train_utils.py` post-licenciement et le pattern de connexion inhabituel (22h-4h) dans `ssh_investigation.md`.
 - Confirmé : Modèle "triche" statistiquement.
 - Preuve : Gain suspect de 93.1% dans `performance_improvements_2024.md` corrélé à la fuite temporelle détectée dans le ticket DATA-889.
 - Confirmé : Violation Éthique.
 - Preuve : Rejet du commit de Thomas pour violation de fairness (Cas Marie) et rapports de `fairness_policy.md` montrant un impact disparate.
 - Écarté : Problème d'infrastructure système.
 - Preuve : `tensorflow_2.15_changes.md` ne montre aucune anomalie en dehors des plages standards.
-

IV. Plan de Validation et Procédures (C2.2)

Plan de Validation Explicite

1. Tests comparatifs (A/B) : Comparaison des poids du modèle actuel avec la version pré-migration dataset pour isoler la dérive.
2. Audit de Fairness : Exécution de tests de non-régression éthique sur les segments sensibles (sexe/âge) pour corriger le *false negative rate*.
3. Validation de robustesse : Stress-test sur l'environnement de test (correction du ticket INFRA-223).

Critères d'acceptation

- Suppression de l'arrondi agressif dans `train_utils.py`.
- Taux de documentation des PRs remontant à > 90%.
- Rétablissement de la parité statistique (Fairness) sur les 3 modèles.

Procédure de Rollback Safe

1. Isolation : Suspension immédiate des accès SSH d'Alex et gel des déploiements automatiques.
2. Restauration : Retour forcé (`git reset`) au commit stable précédent le licenciement d'Alex.
3. Audit de sécurité : Changement des credentials de l'environnement de production.

V. Plan de Retournement à 48h (Scénarios)

- Scénario 1 : Urgence (0-6h)
 - Rollback immédiat des modèles. Révocation des accès. Communication interne.
 - Scénario 2 : Transition (6-24h)
 - Nettoyage du nouveau dataset. Audit du code produit par Thomas. Ré-entraînement partiel.
 - Scénario 3 : Stabilisation (24-48h)
 - Validation des métriques éthiques. Déploiement sécurisé. Rapport de post-mortem client.
-

VI. Registre de Décisions et Communication

Décision	Responsable	Deadline
Gel de la production et Rollback Git	Responsable Code	H+2
Audit de sécurité IP 10.0.2.15	Responsable Infra	H+6
Re-formation urgente de Thomas	RH / Senior Dev	H+12

Communication de Crise

- Interne : Transparence sur le sabotage. Renforcement des règles de "Peer Review" (temps > 1h obligatoire).
 - Clients : "Optimisation en cours des protocoles d'équité algorithmique". Mise en avant de la valeur à long terme via le rétablissement de la généralisation.
-

Annexe : Documentation de Recherche (Données brutes conservées)

- Fichiers consultés : [git_commit.log](#), [ssh_investigation.md](#), [train_utils.py](#), [fairness_policy.md](#), etc.
- Observations : Amélioration suspecte de 93.1%, PRs éthiques PR-385/412 risquées, seulement 45% de doc.

- **Finalité : Ce cas illustre comment une amélioration apparente masque une dégradation profonde révélant une gouvernance insuffisante.**