

10 questions spécifiques au sujet de l'IA d'un point de vue d'expert en IA GenAI

D'accord, en tant qu'expert en IA et en GenAI, voici 10 questions spécifiques concernant l'utilisation de l'IA générative dans les solutions GenAISafety, en tenant compte des produits listés :

1. **Architecture et Modèles:** Quelles architectures de LLM sont privilégiées pour chaque solution (e.g., Transformers, RNNs, autres) et comment sont-elles adaptées aux contraintes spécifiques du domaine HSE (taille des données, latence, etc.) ?
2. **Fine-tuning et Transfer Learning:** Comment le fine-tuning et le transfer learning sont-ils utilisés pour adapter les LLM pré-entraînés aux données spécifiques de chaque client (rapports d'incidents, évaluations des risques, etc.), tout en minimisant le risque de "catastrophic forgetting" ?
3. **Robustesse et Explicabilité:** Quelles techniques sont employées pour garantir la robustesse des LLM face aux données bruitées ou incomplètes, et comment l'explicabilité des décisions prises par les LLM est-elle assurée pour gagner la confiance des utilisateurs ?
4. **Multi-modalité et Fusion de Données:** Comment les LLM sont-ils intégrés avec d'autres modalités de données (images, vidéos, capteurs IoT) pour une analyse plus riche et une prise de décision plus éclairée, et quelles techniques de fusion de données sont utilisées ?
5. **Contrôle et Sécurité des LLM :** Comment les LLM sont-ils contrôlés pour éviter la génération de contenu inapproprié, biaisé ou dangereux, et quelles mesures de sécurité sont mises en place pour protéger les LLM contre les attaques par "prompt injection" ou autres vulnérabilités ?
6. **Adaptation aux Changements Réglementaires:** Comment les LLM sont-ils adaptés rapidement aux changements réglementaires en matière de HSE, et comment l'automatisation de la génération de rapports de conformité est-elle réalisée ?
7. **Évaluation et Métriques:** Quelles métriques spécifiques au domaine HSE sont utilisées pour évaluer les performances des LLM (précision, rappel, F1-score, etc.) et comment ces métriques sont-elles reliées à des résultats concrets en matière de sécurité ?
8. **Edge Computing et Latence:** Comment les LLM sont-ils déployés sur des dispositifs "edge" (e.g., caméras de surveillance, robots d'inspection) pour minimiser la latence et permettre une prise de décision en temps réel, et quelles techniques de compression et d'optimisation sont utilisées ?
9. **Apprentissage Continu et Feedback Humain:** Comment les LLM sont-ils mis à jour en continu avec de nouvelles données et du feedback humain pour améliorer leur précision et leur pertinence, et comment les experts HSE sont-ils impliqués dans le processus d'apprentissage ?

10. **Éthique et Responsabilité:** Comment les questions éthiques liées à l'utilisation de l'IA générative dans le domaine de la HSE sont-elles abordées (e.g., biais algorithmiques, impacts sur l'emploi), et comment la responsabilité des décisions prises par les LLM est-elle définie ?

