

Couplage d'Approches LLM et SLM pour le Déploiement de Solutions d'Extraction d'Entités Nommées

Samuel Kierszbaum (Airbus Protect), Nicolas Heulot (IRT SystemX)*



- Contexte projet
- Approches d'extraction d'entités nommées
- Expérimentation & Résultats

Sémantiques Métier pour l'exploitation de Données multi-sources (SMD)



Kick off

Lancé en juillet 2021 pour 4 ans



**Intelligence
artificielle
et ingénierie
augmentée**

Le troisième projet du programme de recherche « Intelligence Artificielle et Ingénierie Augmentée » (IA2)



**Partenaires
industriels**

Airbus Defence and Space Secure Land communications • Airbus Protect • Ecosys
• EDF • RTE



**Partenaires
académiques**

CentraleSupélec (MICS) • LISN



**Encadrements
académiques**

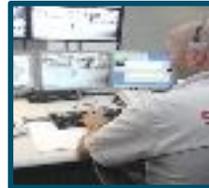
1 thèse, 1 postdoc, 3 stages

Sémantiques Métier pour l'exploitation de Données multi-sources (SMD)

1



Détection et caractérisation de scènes atypiques dans les vidéos de surveillance pour la gestion de crise



2



Aide à la refonte des référentiels métiers dans le contexte des installations nucléaires



3

Rte

Aide à la recommandation dans les salles de supervision du réseau électrique

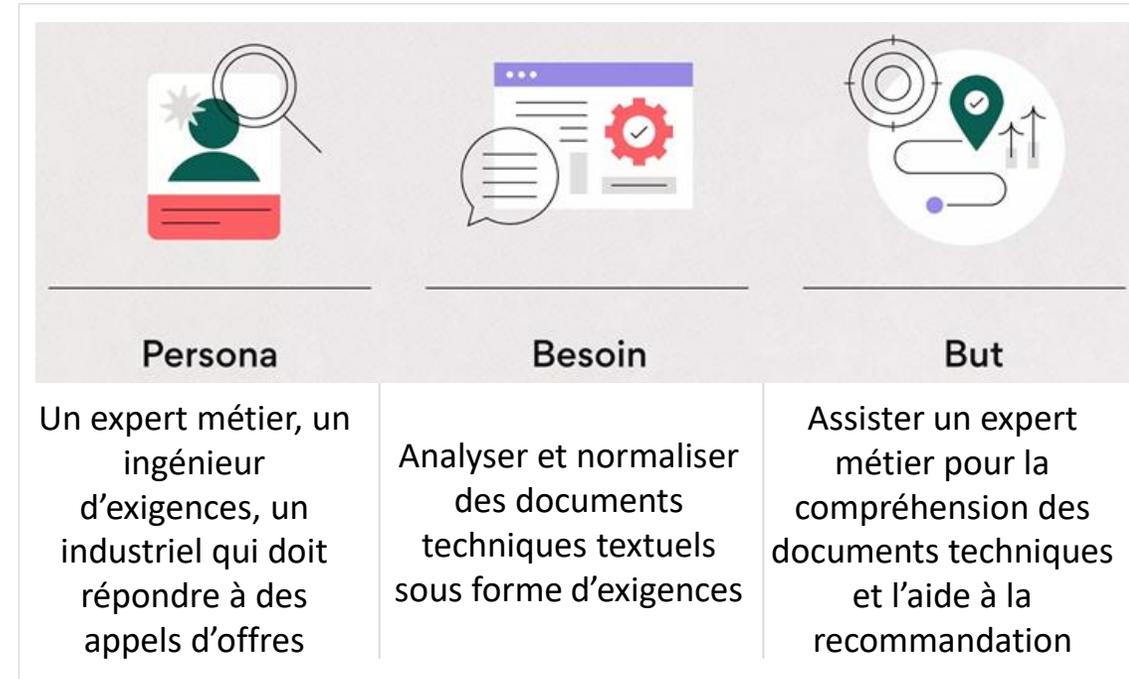
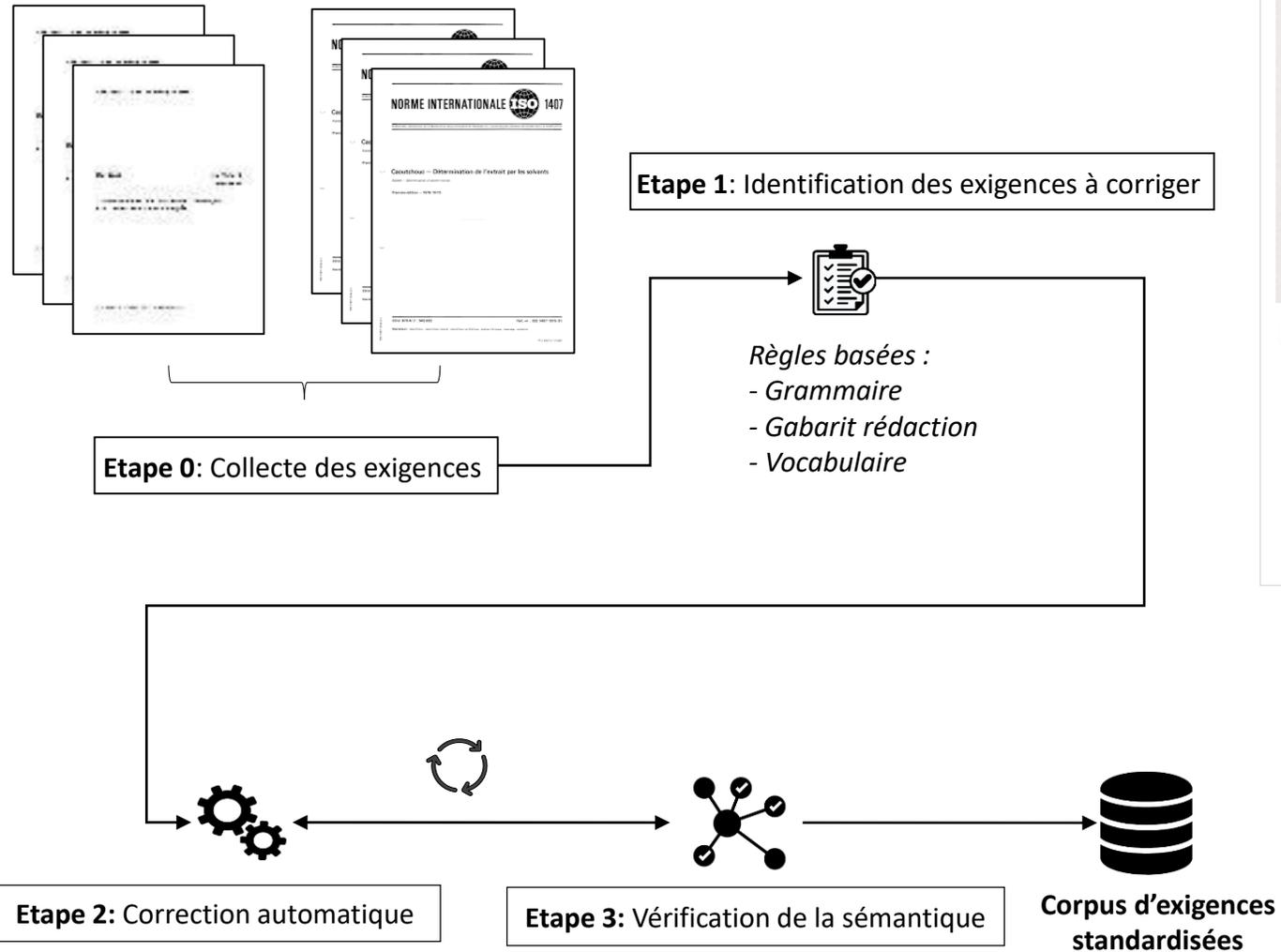


4

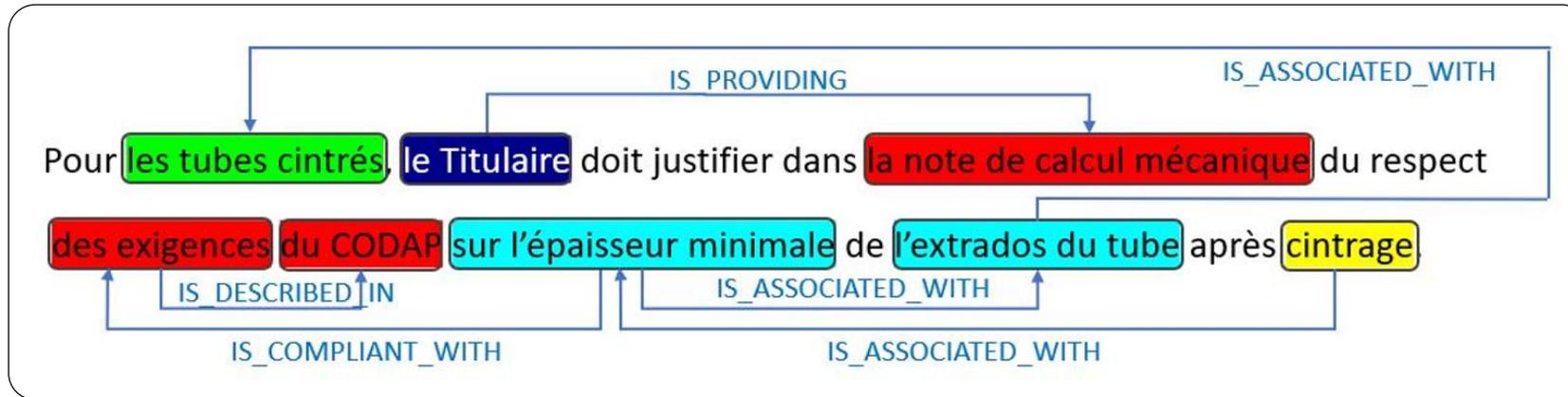


Aide à la recommandation pour les métiers de conseil





- Objectif 1 : Normalisation automatisée (batch)
- Objectif 2 : Recommandations de tournures (interactif)
- Objectif 3 : Propositions de termes (interactif)



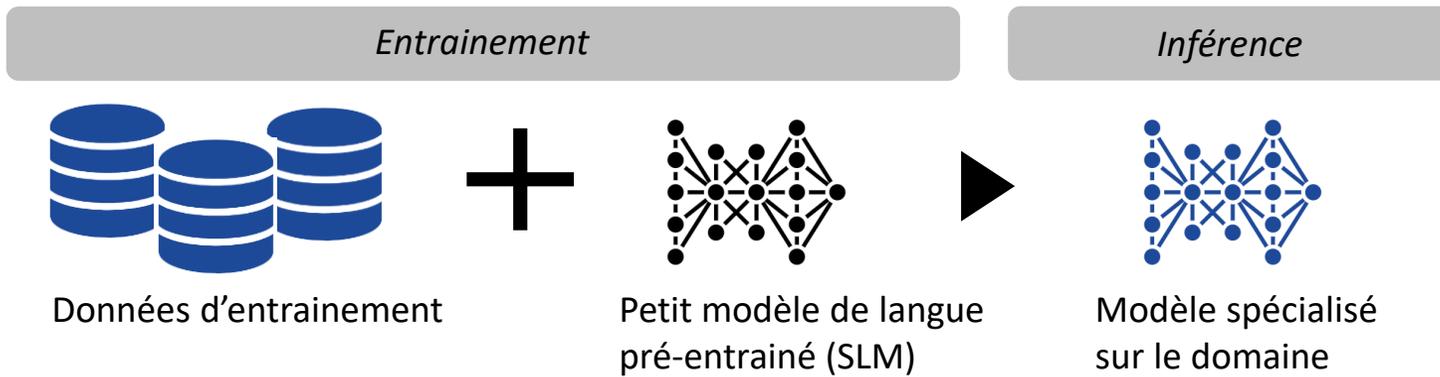
Entités nommées



- L'extraction d'entités nommées est une étape clé pour l'utilisation de l'IA sur des problématiques d'ingénierie des exigences

- ☐ Contexte projet
- ☐ Approches d'extraction d'entités nommées
- ☐ Expérimentation & Résultats

Deep
Learning



Deep
Learning



Données brutes



Guide
d'annotation



Experts



Outil
d'annotation



Données d'entraînement

Inconvénients

- Annotation coûteuse réduisant la taille des données d'entraînement
- Manque de consistance dans les annotations impacte les performances

IA
Généralive

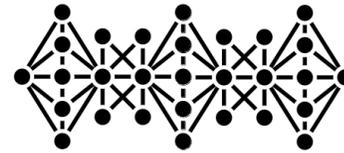
Prompt engineering



Prompt



Inférence

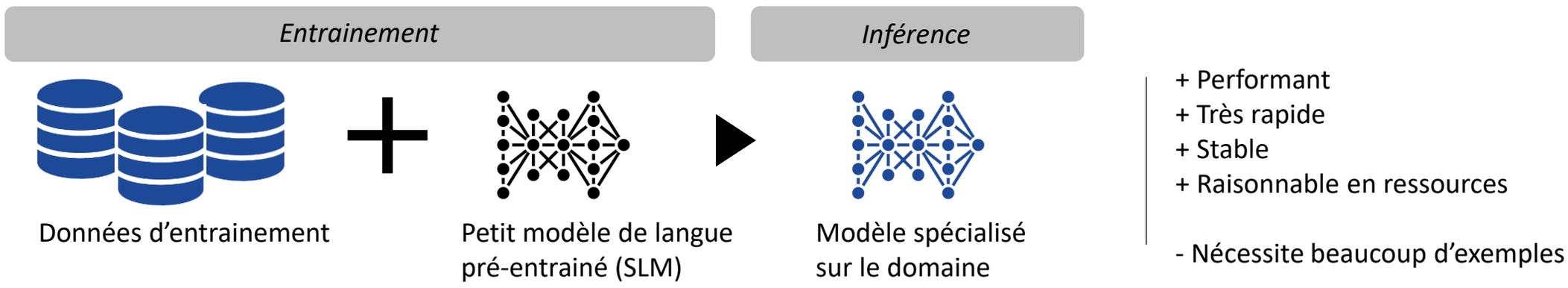


Très grand modèle
de langue (LLM)

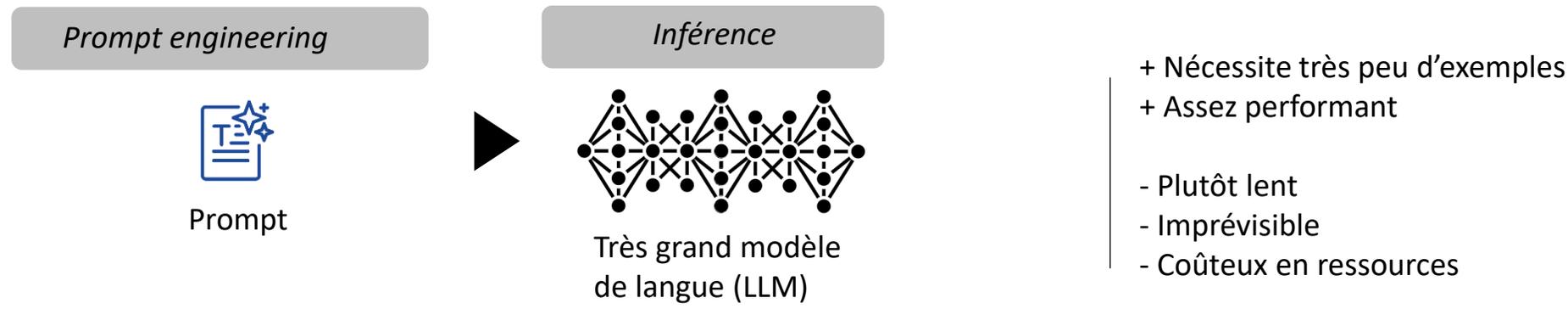
Inconvénients

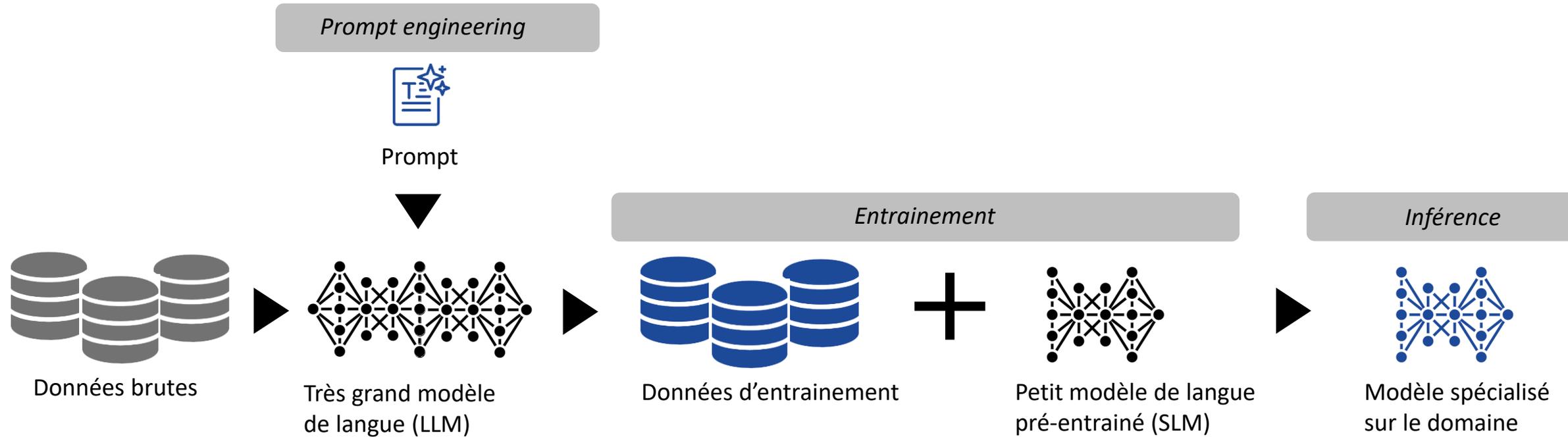
- Travail de prompt engineering « à façon » par entité
 - choix d'extraction mono ou multi entité
 - stratégie de sélection d'exemples
 - système de formatage de la sortie
- Hallucinations des modèles

Deep Learning



IA Générative





- Contexte projet
- Approches d'extraction d'entités nommées
- Expérimentation & Résultats



Objectif : Comparer les approches sur la tâche de reconnaissance d'entités nommées dans des contextes spécialisés avec peu de données annotées manuellement

Baseline: approche basée deep learning, avec fine-tuning d'un SLM sur environ 1000 phrases annotées manuellement

Approche 1: approche basée LLM, few-shot NER, dans un contexte à faibles ressources avec 30 phrases annotées manuellement

Approche 2: approche hybride, avec une étape few-shot NER basé LLM utilisée pour générer des données annotées, qui sont ensuite utilisées pour le fine-tuning d'un SLM



AeroBERT-NER

<https://huggingface.co/datasets/archanatikayatray/aeroBERT-NER>

Tikayat Ray, Archana, Pinon-Fischer, O. J., Mavris, D. N., White, R. T., & Cole, B. F. (2023). aeroBERT-NER: Named-entity recognition for aerospace requirements engineering using BERT. In AIAA SciTech 2023 Forum (p. 2583)

- 1 432 phrases en **Anglais** issues du domaine de l'ingénierie des exigences en aérospatiale
- Annotation en entités nommées selon le schéma d'étiquetage BIO
- 5 entités nommées

SYS	systèmes et matériels
ORG	entreprises et organisations
RES	ressources documentaires
DATETIME	expressions de date et d'heure
VALUE	valeurs numériques

Before December 1, 2012, an electrical or electronic system that performs a function whose failure would prevent the continued safe flight and landing of an airplane may be designed and installed without meeting the provisions of paragraph (a) provided the system has previously been shown to comply with special conditions for HIRF, prescribed under Section 21-16, issued before December 1, 2007.

- Nous considérons des approches de reconnaissance mono-entité:
 - entités SYS (très spécifique au domaine)
 - 999 phrases avec au moins 1 occurrence, pour total de 1855 annotations



Approche	F1 Score	Exact	Partial	Missing	Spurious
fine-tuning	0.90	1563	179	113	238
LLM	0.83	1362	244	249	242
Hybride	0.85	1408	256	191	332

BERT



GPT-4

- Baseline « fine-tuning » avec de bonnes performances
- Faibles écarts entre les approches

- ❑ Résultats encourageants pour l'intégration des LLMs dans l'activité d'annotation afin d'accélérer le déploiement à l'échelle de solutions d'extraction d'entités nommées sur des domaines spécifiques
- ❑ Portée de l'étude à nuancer (choix de modèle, complexité des entités, généralisation de la pipeline)

Perspectives

- Benchmarker avec d'autres outils et modèles (comme flair ou GLiNER)
- Réaliser une analyse de sensibilité sur l'approche hybride afin d'évaluer la robustesse selon la qualité des annotations
- Évaluer le gain en productivité sur le travail d'annotation avec des outils embarquant des fonctionnalités de prompt engineering
- Étendre cette démarche à d'autres tâches d'extraction d'information afin d'utiliser davantage l'IA sur des problématiques d'ingénierie des exigences

Merci pour votre attention

nicolas.heulot@irt-systemx.fr