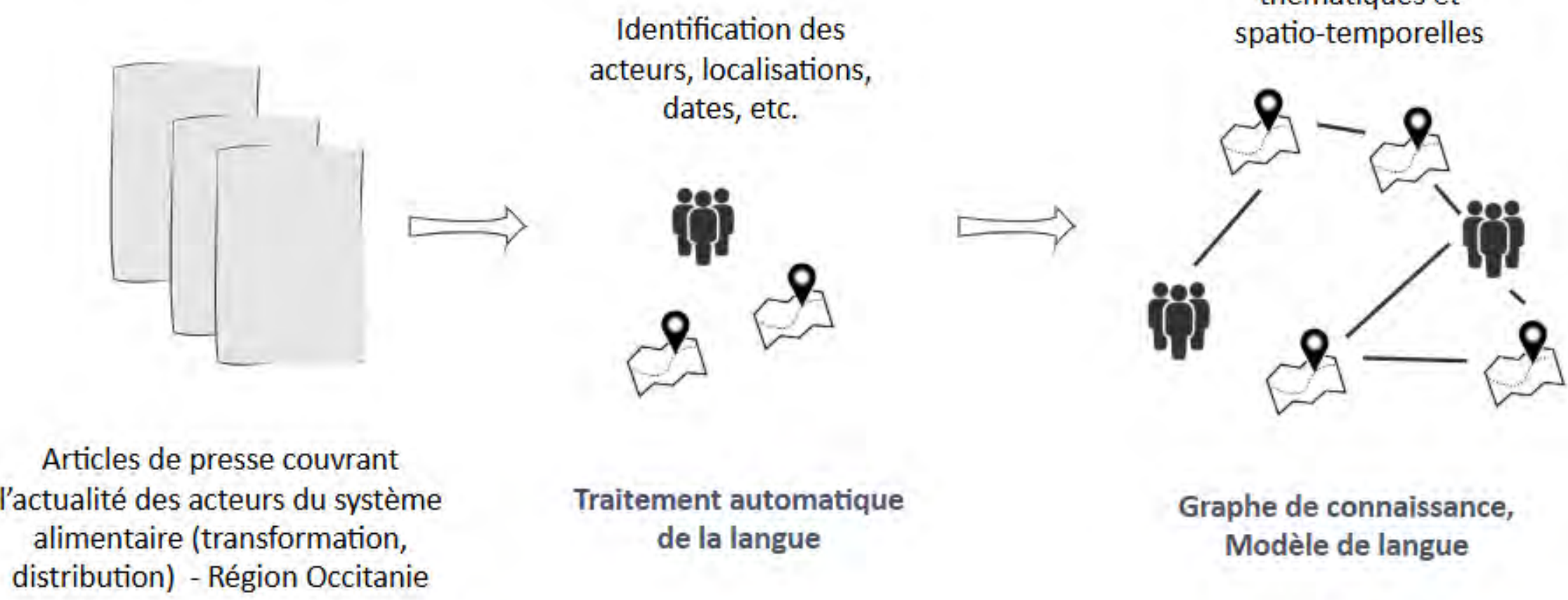


CONTEXTE & PROBLÉMATIQUE



OBJECTIFS & DÉMARCHE

Enjeux méthodologiques



PIPELINE MÉTHODOLOGIQUE

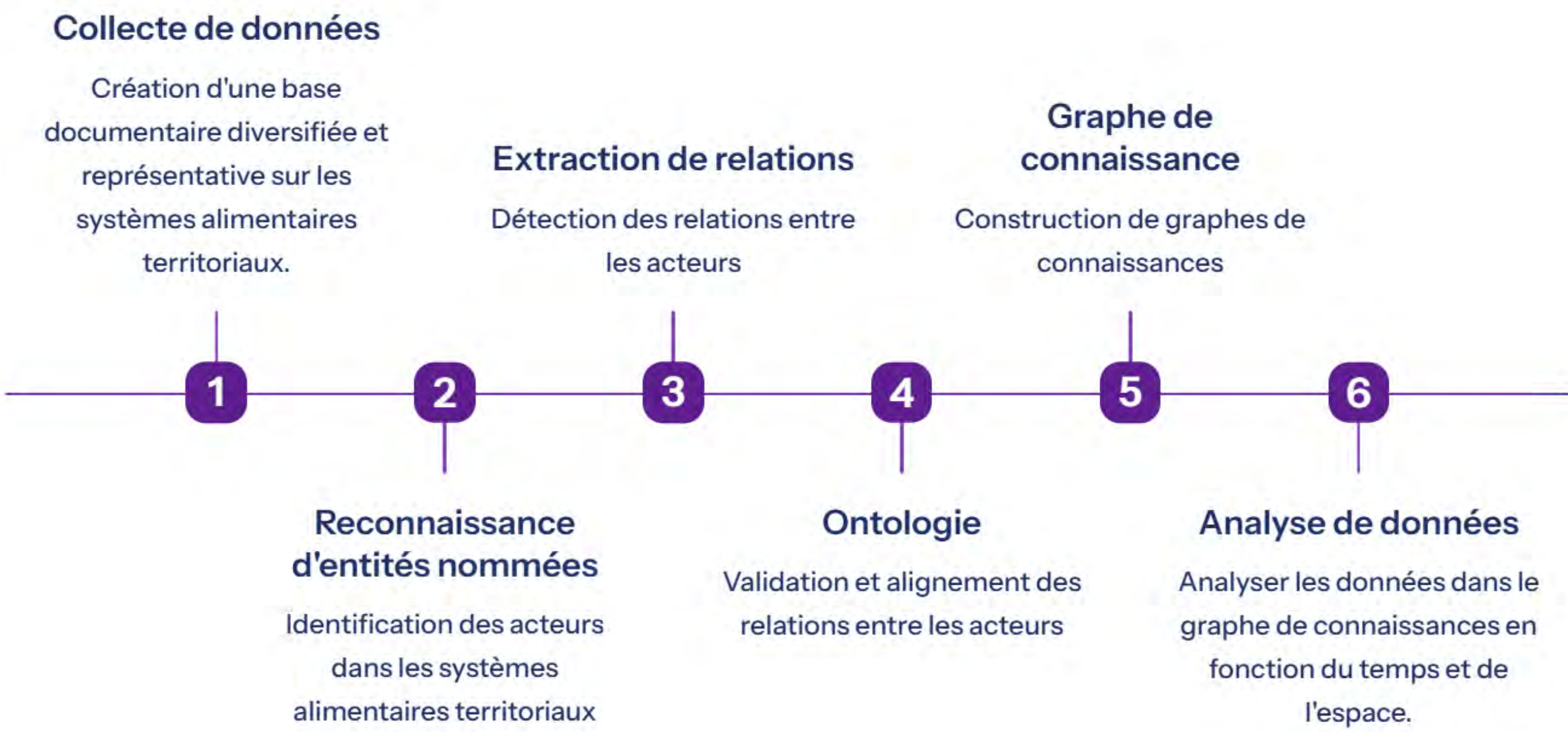
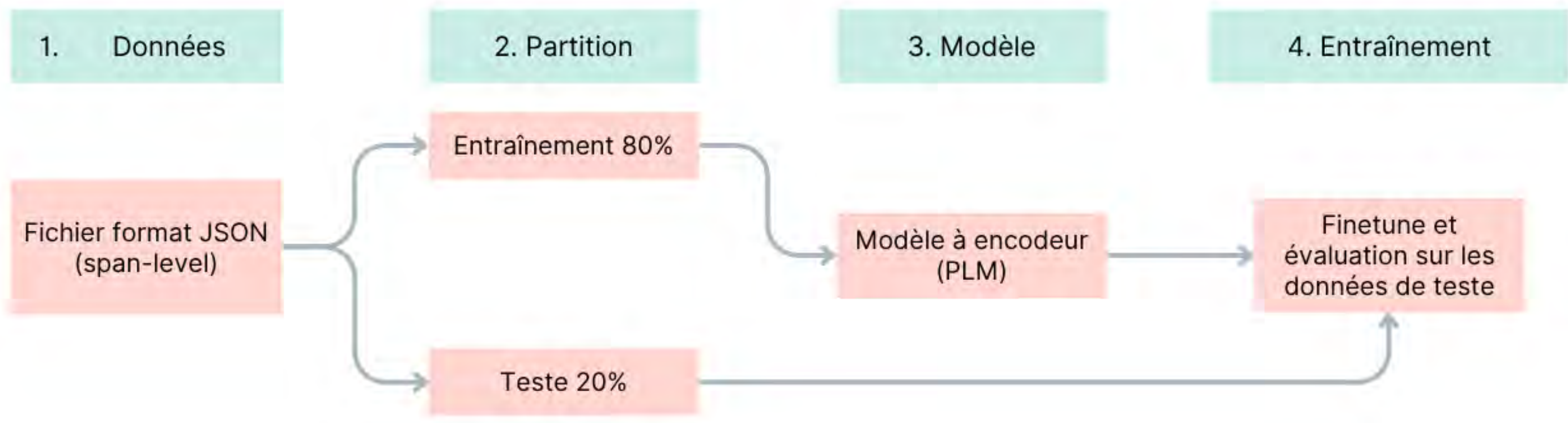


Figure 1. Pipeline de transformation texte → graphe de connaissances

ÉTAPE CLÉ : RECONNAISSANCE D'ENTITÉS (NER)

- Acteurs : producteurs, magasins de producteurs, réseaux
- Lieux : régions, villes, Départements, territoires
- Temporalités : années, périodes, événements

Approche proposée :



RÉSULTATS (NER)

Résultats intermédiaires :

- Corpus annoté 237 articles
- GLiNER [1] (modèle état de l’art).

Approche	F1-Score	Gain
Zero-shot	41,48%	-
Après adaptation	73,14%	+31,66%

Table 1. Performances du NER avant/après adaptation domaine

PROCESSUS D'ANNOTATION (NER étape 2 du pipeline)

- Mis en place d’un guide d’annotation avec les experts en systèmes alimentaires territoriaux
- Organisation d’une session d’annotation collective
- Participation de 14 annotateurs indépendants pour une partie des données
- Revue qualitative avec experts en systèmes alimentaires territoriaux

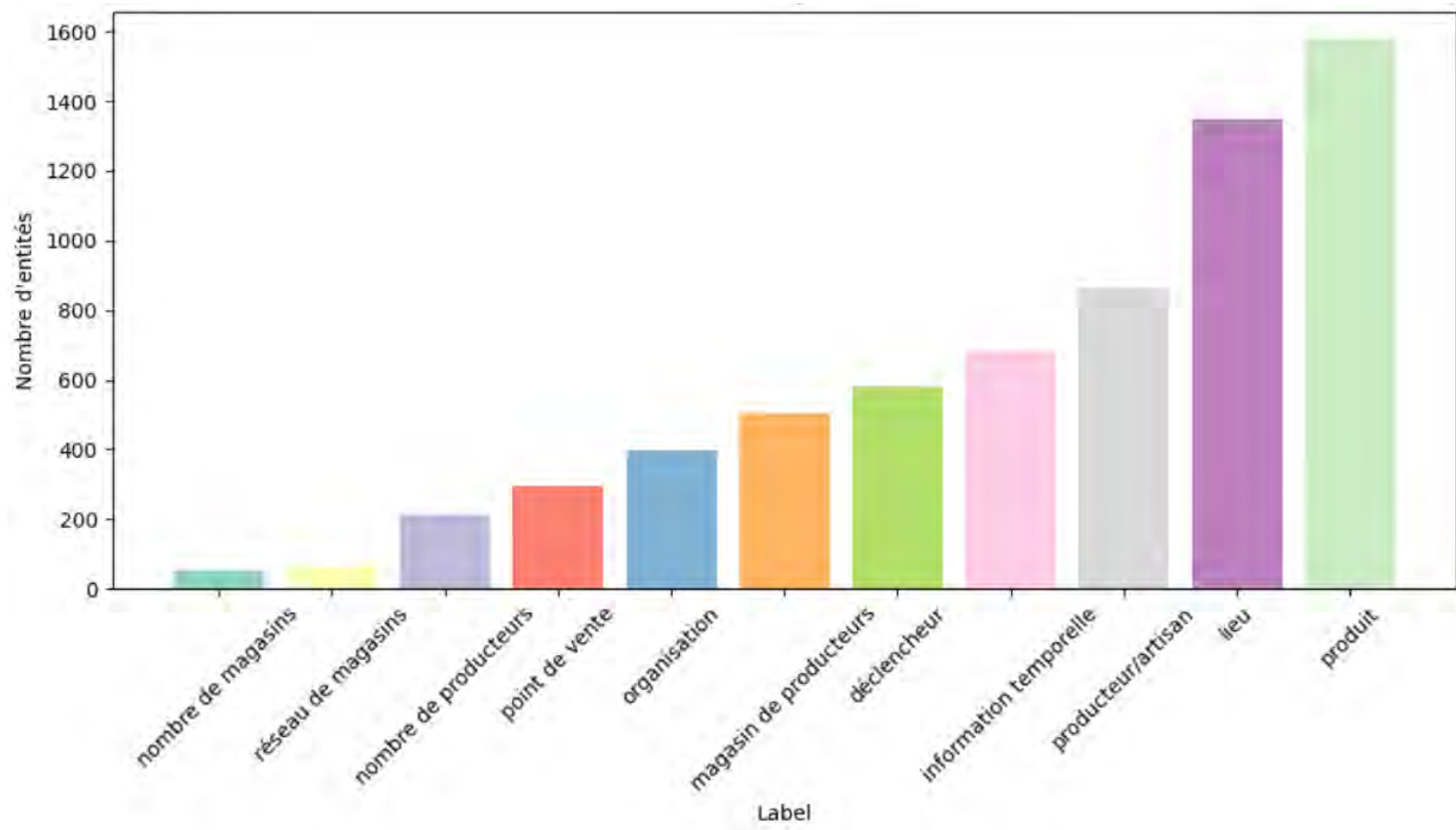


Figure 2. Distribution des entités pour chaque type d'entité sur le corpus total annoté (237 articles)

Extraction de relation (ER) et Ontologie

Construction du corpus

- construction de l’ontologie
- Intégration de l’ontologie dans un LLM pour la validation des relations
- Construction d’un début de corpus pour l’ER

OUTILS & TECHNOLOGIES

- GLiNER : NER zero-shot multilingue
- HuggingFace Transformers : modèles LLM
- Protégé : modélisation de l’ontologie SAT

PERSPECTIVES & IMPACT

Prochaines étapes :

- Construction de corpus complet pour l’extraction de relation et construction automatique du graphe
- Analyse spatio-temporelle des dynamiques locales
- Visualisation interactive pour les collectivités
- Extension à d’autres corpus et territoires

Impacts attendus :

- Politiques publiques : soutien aux stratégies alimentaires territoriales
- Recherche : détection d’acteurs émergents et pratiques innovantes
- Reproductibilité : extension à d’autres systèmes comme la santé ou autres

RÉFÉRENCES CLÉS

1. Zaratiana, U., Tomeh, N., Holat, P., amp; Charnois, T. (2023). Gliner: Generalist model for named entity recognition using bidirectional transformer. arXiv preprint arXiv:2311.08526.

2. Gonzalez De Molina, M., amp; Lopez-Garcia, D. (2021). Principles for designing Agroecology-based Local (territorial) Agri-food Systems: A critical revision. Agroecology and Sustainable Food Systems, 45(7), 1050-1082.

3. Saattrup Nielsen, Dan, Kenneth Enevoldsen, and Peter Schneider-Kamp. "Encoder vs Decoder: Comparative Analysis of Encoder and Decoder Language Models on Multilingual NLU Tasks." arXiv e-prints (2024): arXiv:2406.

4. Bogdanov, S., Constantin, A., Bernard, T., Crabbé, B., amp; Bernard, E. (2024). NuNER: entity recognition encoder pre-training via LLM-annotated data. arXiv preprint arXiv:2402.15343.

5. Zhou, W., Zhang, S., Gu, Y., Chen, M., amp; Poon, H. (2023). Universalner: Targeted distillation from large language models for open named entity recognition. arXiv preprint arXiv:2308.03279.