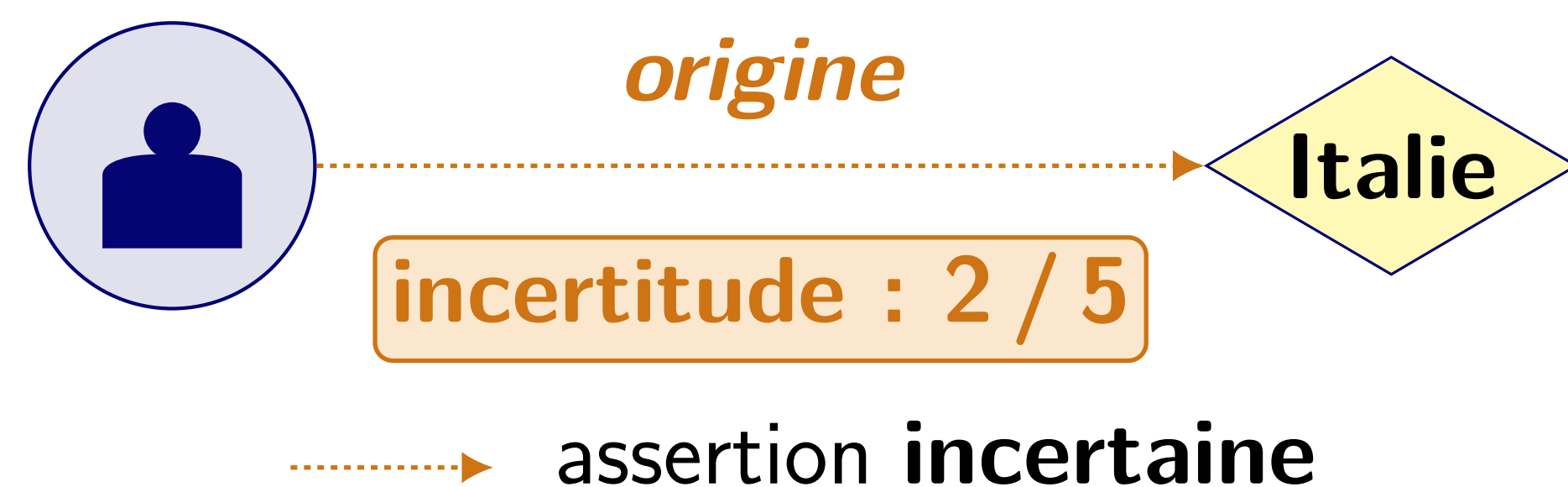


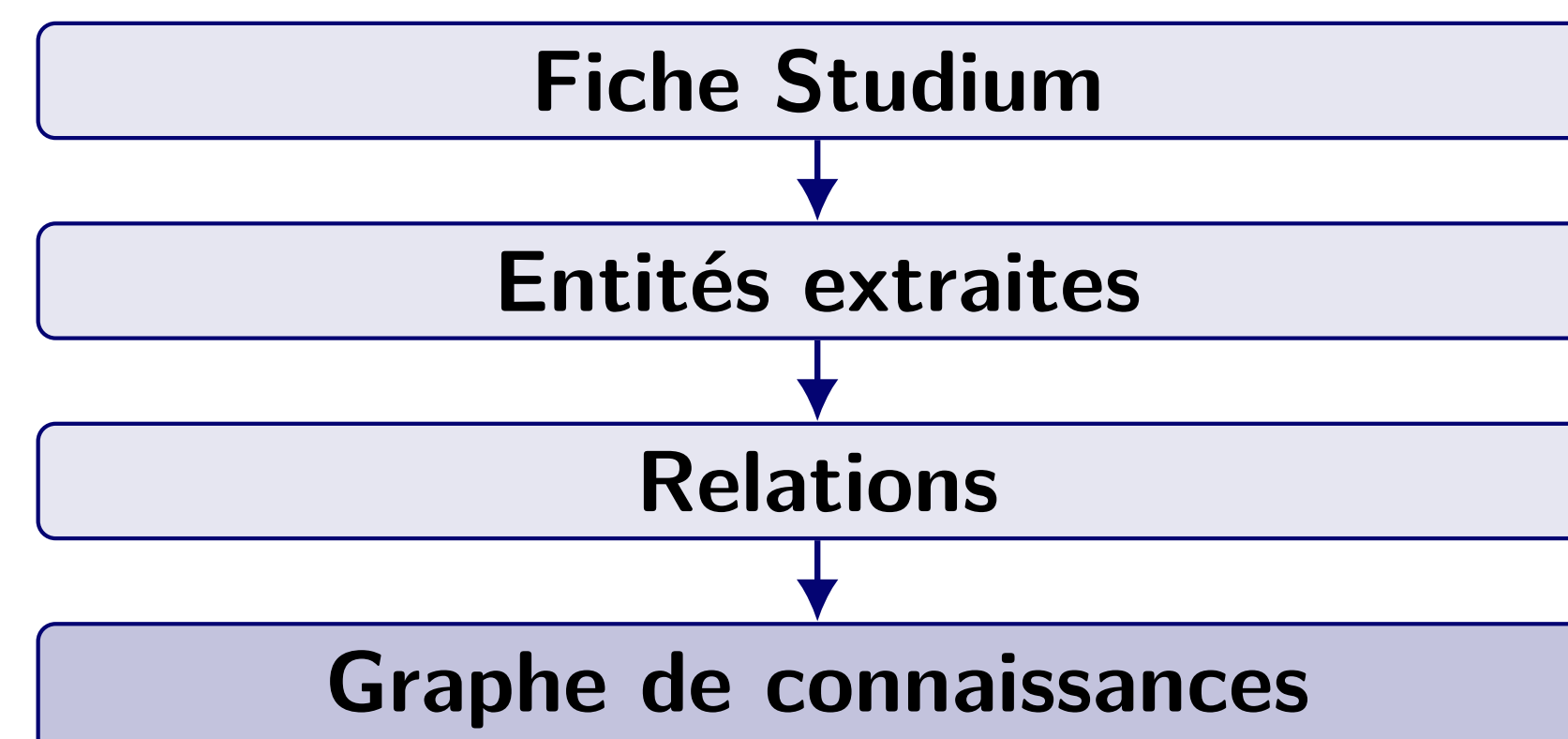
## Introduction

**Problème** : Les fiches historiques contiennent des informations **incertaines**, **imprécises** ou **manquantes**.

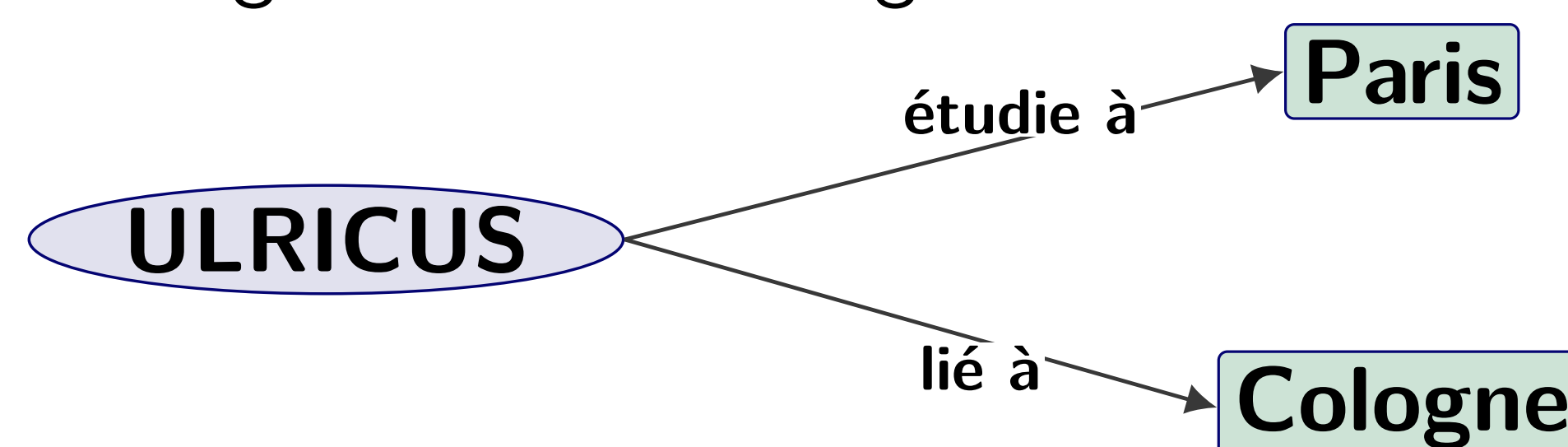
**Objectif** : Analyser ces informations en vue de leur intégration dans un **graphe de connaissances**.



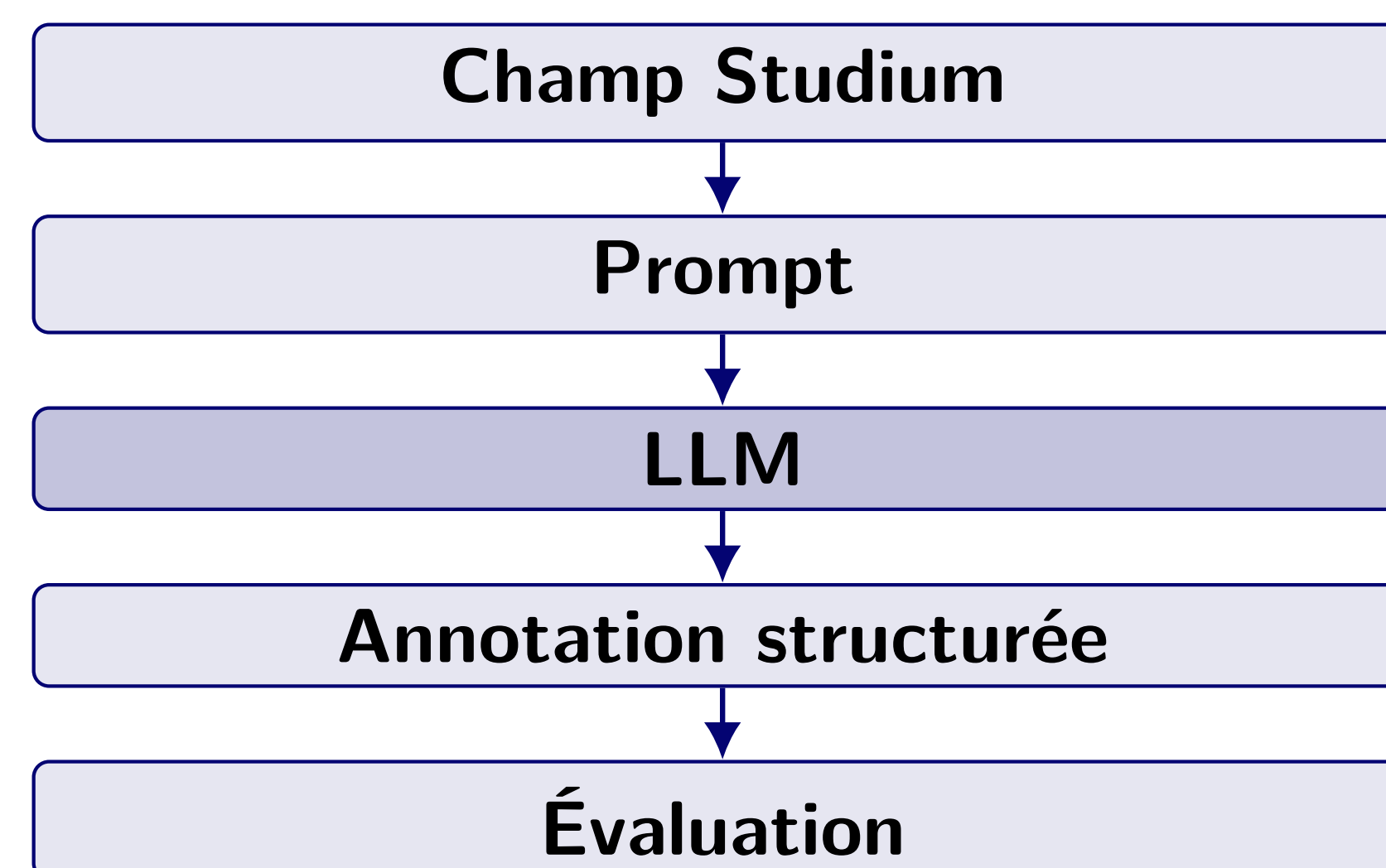
## Du texte au graphe



**Entités** : personne · lieu · institution · date  
**Relations** : origine · affiliation · grade · relation · source



## Pipeline



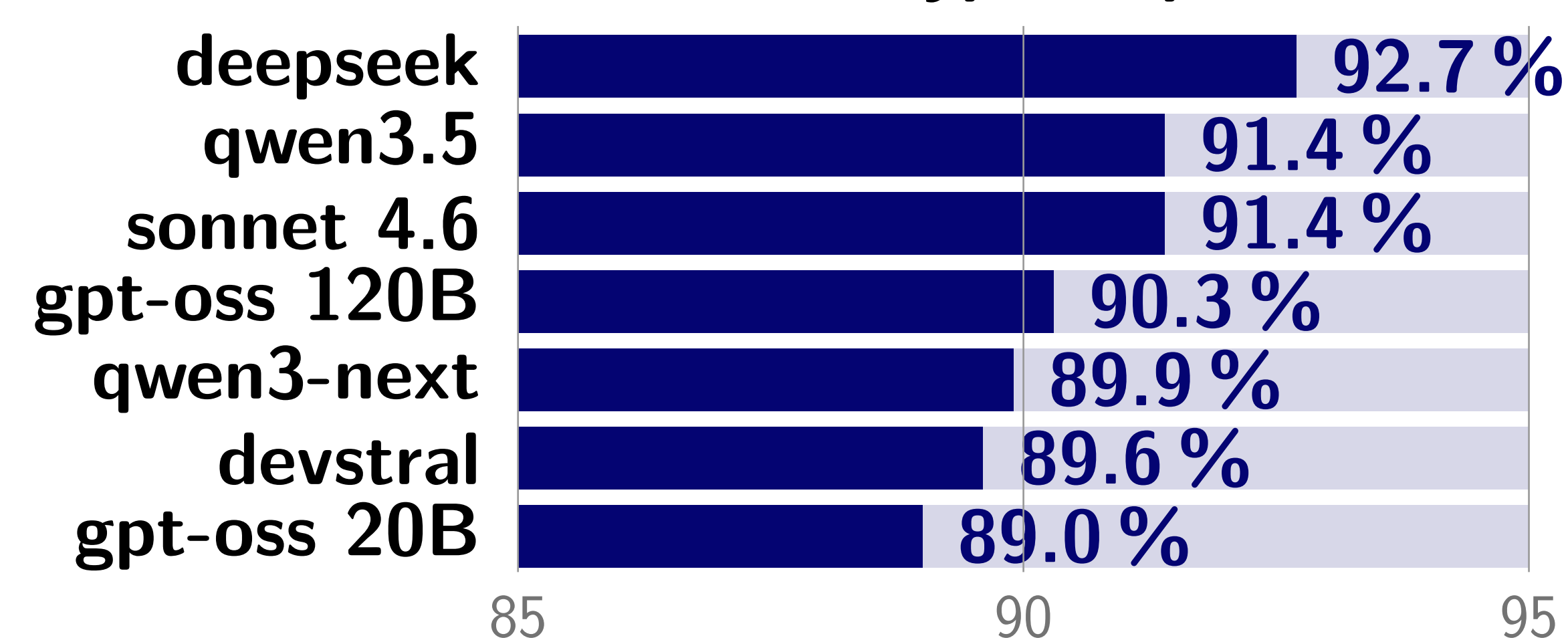
206 exemples    158 positifs    48 négatifs

**Évaluation** : type · portée · certitude

## Résultats

7 LLM testés sur 206 exemples. **Détection binaire** (incertitude, imprécision, manquant, aucune) : **95-99 %**.

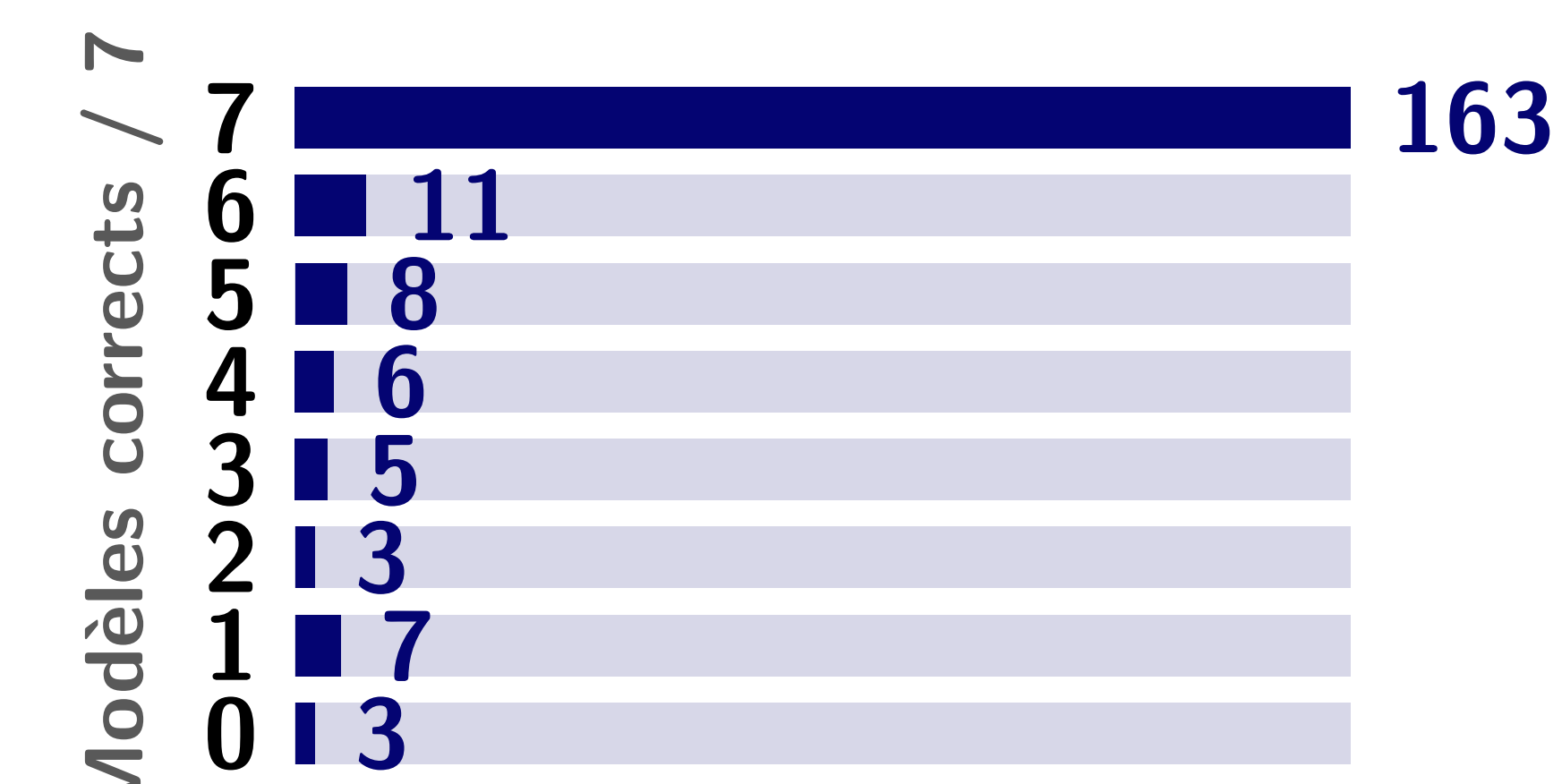
**F1 sur annotation : type + portée**



## Difficulté des exemples

**Distribution des performances des modèles (206 ex)**

(annotation correcte = type + portée)



Nombre d'exemples

**79 %**

exemples réussis par les 7 LLM testés

## Cas limites

**Paris : lieu ou institution ?**

Cursus ▶ Université : Paris (Nation de France) ?

**Paris = ville**  
→ lieu

**Paris = Université**  
→ institution

**Annotation de référence = institution**  
Des modèles répondent **lieu**, d'autres **institution**.

« Il a également séjourné auprès de son cousin à Avignon »

**Annotation possible :**

- ▶ cousin comme **entité** → information manquante
- ▶ **relation** familiale seule → information suffisante

**La granularité du graphe change l'annotation.**

## Perspectives

- ▶ Enrichir le benchmark
- ▶ Analyser les **erreurs de portée**
- ▶ Tester d'autres **prompts** et d'autres **modèles**
- ▶ Intégrer ces annotations dans le **graphe**



## Exemples de données Studium

**Variantes de nom** : Ulrich de STRASBOURG

Uldaricus ENGELBERTI

**Lieu de naissance** : **Italie ?**    **Date de vie** : -1278

**Université** : ?

**Activité** : **vers 1248-1254**    **Grade** : Bachelier en théologie (Paris)

**Commentaire** : « Son parent est **peut-être** Guillelmus de VISSACO, sacristain de Cluny »

**Commentaire** : « Il est mort à Paris avant d'avoir complété son cursus pour la maîtrise **?** »

**Identité** : « L'Adam de KALKER n'est **peut-être qu'un fantôme** »

incertitude    imprécision    variante / manquant

Extraits de fiches différentes