Sommaire

- 1. Description de la tâche
 - 1.1. Entités nommées sur INCEpTION
 - 1.2. Protocole
- 2. Retour d'expérience

Mots-clés: entité nommé, couche d'annotation, span

1. Description de la tâche

Une entité nommée est une expression linguistique qui désigne soit :

- une personne (réelle, fictive, historique, contemporaine, etc.)
- un lieu (ville, pays, nom de lieu, etc.)
- une organisation (entreprise, institution, société, etc.)

La tâche consiste donc à rechercher les mots et/ou groupes de mots appartenant à l'une de ces catégories.

1.1 Annotation des entités nommées sur INCEpTION :

L'annotation des entités nommées peut se faire à l'aide d'une couche intégrée à la plateforme INCEpTION (Named Entity). C'est une annotation de type Span (cf Fiche n°7 - Les types de relations), avec des paramètres réglés par défaut. Certains sont modifiables, d'autres non :

- Enable : ce paramètre permet d'autoriser la possibilité de supprimer la couche d'annotation. Par défaut, il est activé, mais il est possible de le désactiver, et donc d'autoriser la suppression de cette couche.
- Read only: ce paramètre permet de choisir si cette couche est éditable ou juste visualisable. Par défaut, ce paramètre est désactivé mais il est possible de l'activer pour que la couche soit éditable.
- Validation: ce paramètre permet d'activer ou de désactiver la vérification automatique de la bonne formation des annotations (selon les autres paramètres pré-définis) au moment de la clôture du document. Par défaut, la vérification se fait à chaque fois (Always), mais il est possible de ne jamais la faire automatiquement (Never).

- Granularity: ce paramètre définit le niveau sur lequel l'annotation porte. Le paramètre définit est "Tokens-level", c'est-à-dire sur plusieurs mots. Ce paramètre est modifiable.
- Overlap: ce paramètre permet ou interdit la superposition et l'enchâssement des étiquettes. Ce paramètre est réglé sur "None", rendant impossible la superposition et l'enchâssement. Il est modifiable.
- Allow crossing sentence boundaries : ce paramètre est désactivé par défaut mais il est modifiable. Il permet le franchissement des limites de la phrase.

Pour plus de détails sur les paramètres d'une couche d'annotation, voir Fiche n°8 - Création d'une couche d'annotation.

1.2 Protocole:

Étape 1: Choisir la couche d'annotation Named entity



Image 1: Couche d'annotation Named Entity

Étape 2 : Surligner les entités nommées



Image 2 : Sélection des entités nommées

Étape 3 : Aller dans value et choisir la catégorie adéquate

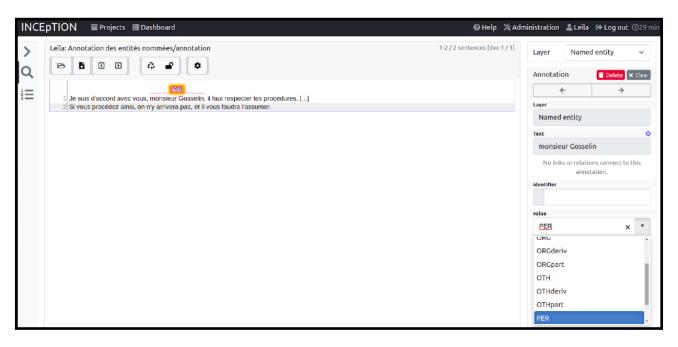


Image 3 : Sélection de la catégorie d'entité nommée

Jeu d'étiquette pour les entités nommées disponible sur INCEpTION :

• PER: personne

• LOC: lieu

• ORG: organisation

• OTH: autre

- PER/LOC/ORG/OTH + deriv : mot dérivé d'une entité nommée (ex: la ville rose → Toulouse)
- PER/LOC/ORG/OTH + part : pour un mot contenant un trait d'union, la partie surlignée appartient à la catégorie renseignée.

2. Retour sur expérience

Avec INCEPTION, l'annotation des entités nommées est intuitive et ne pose pas de problème en particulier.