



DGI – DTI
Unité Opérations et
support aux utilisateurs

Demande d'installation

Formulaire de demande de validation d'un besoin informatique

Version 0.8
Septembre 2018

Nom du demandeur	JAGER
Login	JAGERBE
Service	DTI/CSI
Tél. / Bureau	6141 /
Date de la demande	9/2/2023
Numéro de la machine	N/A (Licence flottante)

Description du besoin :

Afin d'améliorer la pertinence de la détection d'éléments dans les textes de la CJUE, tout autant que de mettre en place de nouvelles détections, nous avons besoin d'une application permettant d'annoter les textes. Prodigy convient parfaitement à ce besoin (<https://prodi.gy/buy/>), pour un coût de 2450 USD.

Justification de la demande (description des tâches à accomplir, nom du projet) :

Toute demande non justifiée sera automatiquement refusée

L'annotation de textes est fondamentale pour les projets de détection automatique dans des textes. Il s'agit de surligner les éléments qui nous intéressent (par exemple des références juridiques), de les typer de diverses façons (par exemple références directes ou indirectes), et d'exprimer des liens de natures diverses entre elles (par exemple lien entre une référence indirecte et une directe comme dans "article 40 TFUE..... l'article 25 du même traité" qui lie une référence indirecte avec une référence directe précédente).

Tous les projets d'IA s'appuient sur des données pour former les modèles de prédiction. Le pipeline de formation comporte généralement une étape d'annotation pour les données non structurées (comme le texte, les images, la parole, etc.). Les meilleures pratiques pour l'étape d'annotation recommandent d'utiliser un outil spécialisé plutôt que Word, Excel ou d'autres outils d'édition de documents. Prodigy de Spacy.io permet aux utilisateurs d'annoter à partir de zéro ou de corriger des annotations, pour différents cas d'utilisation, dans un environnement collaboratif qui optimise le processus d'annotation. Prodigy est un outil d'annotation spécialisé traitant un large éventail de tâches d'apprentissage automatique, y compris la reconnaissance d'entités nommées, la catégorisation de portée, la classification de texte, le balisage des parties du discours, l'analyse des dépendances, la résolution de coréférence, la vision par ordinateur, l'annotation audio et vidéo, la transcription, le diarisation de locuteur, formation et évaluation, etc. L'outil peut importer des

annotations existantes et il peut exporter celles-ci en utilisant des formats standardisés, faciles à intégrer dans un pipeline d'apprentissage de l'IA

Durée du projet :

L'utilisation de ces licences est envisagée pour tous les futurs projets de l'innovation Lab.

Population (services, nombre de personnes impactées) :

5 personnes

Nom du supérieur hiérarchique	Eric DEUDON
Visa	
Date	17/04/23

Formulaire à retourner complété au Helpdesk, scanné par mail (Hdb.Help_Desk_Informatique@curia.europa.eu) ou par courrier au (TM/00LB0479).

Cadre réservé à la Direction des Technologies de l'Information

Risques de sécurité	Inexistants <input type="checkbox"/>	Faibles <input type="checkbox"/>	Forts <input type="checkbox"/>	Critiques <input type="checkbox"/>
Nécessité de droits admin ?	Oui <input type="checkbox"/>	Non <input type="checkbox"/>		
Décision finale d'installation	Oui <input type="checkbox"/>	Non <input type="checkbox"/>		
N° de call Help Desk				