

Article preprocessing

Format : image docker, programme python

Rôle

Responsable du bon déroulement de l'étape de preprocessing des articles qui consiste pour chaque article d'épurer des mots inutiles et de construire leur tf-idf.

Flux de sortie : Envoie des tf-idf de chaque article vers le module parent
Transfert des articles vers les modules enfants
preprocessing

Flux d'entrée : Réception des tf-idf de chaque article depuis les modules enfants de preprocessing
Réception des articles depuis le modules parents

Api : envoie des articles et réception des tf-idf par api_Transfert
communication avec le module parent par http_requete
creation des modules enfants avec ssh_launch_program

Arguments

mot de passe et coordonné de la machine-performante
coordonné du module parent