

OpenText Named Entity Recognition

Fournit un service de réduction des risques à vos utilisateurs finaux qui cherchent à se conformer aux réglementations et aux directives de l'entreprise, en utilisant l'extraction d'entités nommées et nos grammaires organisées.



Avantages

- **Conformité défendable** grâce aux informations sur les sources de données et aux exemples de types de correspondance
- **Mise sur le marché rapide** avec des grammaires prêtes à l'emploi
- **Large prise en charge réglementaire** couvrant de nombreux pays, langues et types d'entités

Quand bien même les performances optimales sont essentielles pour trouver et extraire des entités sensibles et à forte valeur ajoutée à partir de données structurées, semi-structurées et non structurées, la rapidité du moteur n'est pas le seul facteur important.

Il est nécessaire de prendre en charge les opérations avant et après la correspondance afin d'effectuer des optimisations telles que le contrôle de l'étendue, la validation différée et la terminaison anticipée sélective pour des gains de performances immédiats, ainsi que la composantisation, l'extraction et la normalisation pour plus d'efficacité en aval.

Le moteur OpenText™ Named Entity Recognition est un composant logiciel flexible et intégrable qui recherche et extrait les entités sensibles et à forte valeur ajoutée à partir de données structurées, semi-structurées et non structurées à l'aide d'un traitement du langage naturel hautes performances.

Il est conçu pour fonctionner harmonieusement avec des composants tiers à l'aide d'API et d'options architecturales facilement utilisables. Lorsque la solution est associée à d'autres SDK, tels que OpenText™ File Content Extraction pour l'extraction de texte, la fonctionnalité multicomposants supplémentaire est déverrouillée, comme l'extraction de tableaux, afin d'améliorer la précision, les pipelines, et, par extension, les performances.

Autres avantages clés

Utilisez OpenText Named Entity Recognition avec des grammaires organisées pour fournir un service de réduction des risques à vos utilisateurs finaux qui cherchent à se conformer aux réglementations et aux directives de l'entreprise. Nous fournissons :

- Une différenciation supplémentaire avec des grammaires organisées et des grammaires personnalisables par l'utilisateur final.
- Des décisions explicables grâce à l'utilisation de grammaires organisées à la main plutôt que de modèles dérivés de corpus.
- Une réduction des coûts matériels : la latence et le débit peuvent être contrebalancés par la précision et le rappel, afin de répondre aux objectifs de coûts à n'importe quelle échelle.
- Une réduction des coûts d'ingénierie continus à mesure que les réglementations et les définitions d'entités évoluent. Nous veillons à ce que les grammaires restent à jour et performantes.
- Une équipe expérimentée, avec l'expérience nécessaire pour comprendre comment créer les grammaires les plus efficaces. Avec une bibliothèque de centaines de grammaires disponibles, notre expérience n'est plus à prouver.

Conformité défendable

OpenText Named Entity Recognition est conçue avec des grammaires qui définissent les entités qui vous intéressent. Vous bénéficiez d'un accès complet à l'expertise établie d'OpenText en matière de conception de grammaires, grâce à nos modules hautement ciblés, qui fournissent des paquets optimisés, contextuels et maintenus pour des classes d'entités spécifiques à des secteurs géographiques et à des industries afin de répondre aux cas d'utilisation courants tels que :

- Données personnelles/données personnelles identifiables, y compris la détection rapide et précise des adresses
- Données de santé protégées
- Entités de l'industrie des cartes de paiement

Ces grammaires permettent à votre produit d'offrir des fonctionnalités étendues de conformité et de gouvernance comme cas d'utilisation principal, immédiatement et sans développement de grammaire préalable ni coûts d'ingénierie continus au fur et à mesure de leur organisation.

Mise sur le marché rapide

Votre produit est unique, et OpenText Named Entity Recognition vous donne le choix parmi une variété de formats, tout en prenant en charge plusieurs langages de programmation, pour trouver la solution la mieux adaptée à votre environnement. Vous pouvez choisir d'optimiser la vitesse ou la précision, et dans quelle mesure. Vous pouvez également choisir de maximiser le nombre d'entités découvertes ou de limiter les faux positifs en utilisant le score de confiance renvoyé pour chaque entité découverte.

Combinez ces options pour trouver le juste équilibre entre précision, rappel, performances et commodité, mais surtout l'équilibre qui vous convient le mieux, à vous et à votre produit.

Large prise en charge réglementaire

Avec des grammaires organisées couvrant des entités de données personnelles/données personnelles identifiables (PII) dans 46 pays et 42 langues, ainsi qu'une partie de nombreux pays et langues, conçues pour des réglementations telles que le RGPD (Union européenne), le CCPA (Californie), la KVKK (Turquie), la POPI (Afrique du Sud), les APP (Australie), la NZPA (Nouvelle-Zélande) et la CPPA (Canada). Entités de données de santé protégées (PHI) couvrant les types d'entités et conçues pour les industries réglementées, telles que les entités HIPAA (santé) et PCI (industrie des cartes de paiement).

La solution englobe également les classifications gouvernementales, y compris les marquages DoD, les informations non classifiées contrôlées (CUI), le contrôle des exportations et l'Australian Protective Security Policy Framework.

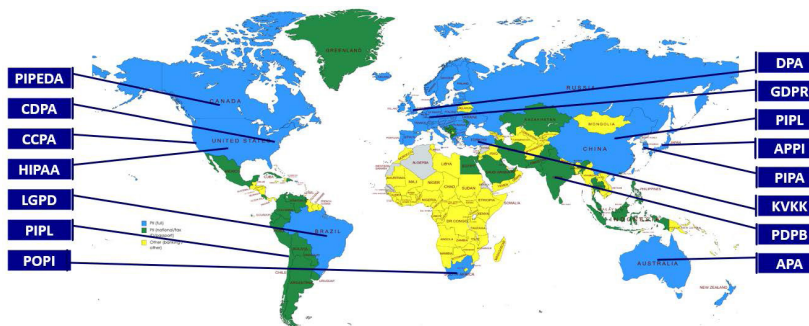


Figure 1 : Régions de la découverte des PII dans le monde

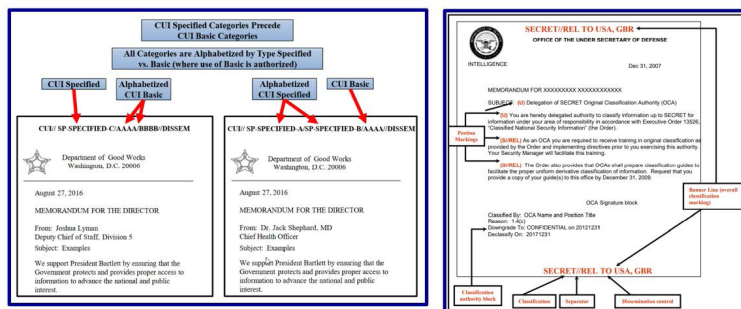


Figure 2 : Exemples de reconnaissance d'entités

Quest Diagnostics est l'un des principaux prestataires de soins de santé aux États-Unis, où il dispose de plus de 2 000 centres ouverts aux patients. Il utilise OpenText Named Entity Recognition dans OpenText Knowledge Discovery pour une indexation, une recherche et une récupération rapides et approfondies.

Quest Diagnostics tire parti de plus de 20 ans de données contenues dans 3,4 milliards de documents.

[En savoir plus >](#)

Fonctions supplémentaires	Description
Prise en charge de grammaires hybrides	Inclut les listes de dictionnaires, les expressions régulières, les grammaires de composants et les combinaisons.
Traitement contextuel	Utilise la structure et les repères pour améliorer le rappel et réduire les faux positifs.
Extensions de grammaire personnalisées	Qui peuvent être mises à la disposition des utilisateurs finaux ou faire partie d'un service à valeur ajoutée.
Grammaires compilées	Pour plus d'efficacité et de protection de la propriété intellectuelle, mais toujours extensibles.
Score de confiance	Permet un contrôle précis de la courbe de précision/rappel.
Marquage des composants	Permet de déterminer précisément quels octets correspondent aux composants d'une grammaire, pour un traitement en aval.
Fonctionnement entièrement diffusé	Pour une utilisation dans les architectures en pipeline (comme avec le composant d'extraction de texte d'OpenText File Content Extraction) pour une latence et une utilisation de la mémoire réduites.
Terminaison anticipée	Réduit l'utilisation du CPU lorsqu'un risque est identifié ou exclu avant le traitement d'un fichier entier.
Optimisations du prétraitement	Contrôle avec précision l'étendue pour limiter l'utilisation du CPU, réduire la latence et augmenter le débit.
Validation des correspondances	S'assure que les codes de détection et de correction des erreurs, tels que les sommes de contrôle, sont valides, réduisant ainsi les faux positifs.
Normalisation	Normalise les résultats renvoyés, tels que les dates, les adresses et les noms, pour un traitement en aval cohérent et un risque réduit.
Gouvernance automatique	Masque, restreint et supprime les entités là où elles se trouvent pour assurer la conformité.
Analyse des sentiments	Fournit une analyse au niveau de la clause, ce qui signifie que chaque sujet d'une phrase reçoit son propre score de sentiment.

Options de déploiement d'OpenText Named Entity Recognition :

Développez votre équipe

- Logiciel sur site, géré par votre entreprise ou OpenText

Ressources

[Demander une démonstration >](#)

[Page Web principale >](#)

[Nouveautés - Blog actualisé en continu >](#)

Configuration système requise par format

Bibliothèque intégrable (SDK Named Entity Recognition)

- Linux (x32 et x64), Windows (x32 et x64), macOS (x64, M1)
- C/C++, .NET, Java

Microservice (serveur Named Entity Recognition)

- Linux (x64), Windows (x64)
- Service basé sur les processus
- API basée sur HTTP
- Nécessite le service de licence Knowledge Discover

Processeur NiFi

- Linux (x64), Windows (x64)
- API NiFi
- Nécessite le service de licence Knowledge Discover

À propos d'OpenText

OpenText, The Information Company, permet aux organisations d'obtenir des informations grâce à des solutions de gestion des informations leaders sur le marché, sur site ou dans le cloud. Pour plus d'informations sur OpenText (NASDAQ : OTEX, TSX : OTEX), consultez : opentext.com.