

SYSTRAN

Linguistic Development Kit (LDK)

Créez les solutions dont vos clients ont toujours rêvé !

Parce que le développement de moteurs de traduction automatique s'appuie sur des technologies de pointe, SYSTRAN demeure à l'avant-garde des technologies de traitement du langage naturel depuis plusieurs décennies.

Disponibles sur l'ensemble des langues au catalogue, SYSTRAN vous propose désormais les services linguistiques suivants :

- Filtrage de document
- Détection de la langue
- Segmentation et tokenization
- Normalisation de la langue
- Catégorisation de documents
- Reconnaissance d'entité nommée
- Dictionnaires
- Analyse morphologique
- Analyse syntaxique
- Translittération
- Résolution des ambiguïtés



Pris séparément, chaque module est un outil très efficace pour traiter des langues, des documents ou des entités nommées. En les combinant, vous obtenez des capacités de traitement multilingues performantes pour vos solutions de text-mining ou d'analyse sémantique.

Quels que soient les besoins de vos clients finaux (eDiscovery, procédure judiciaire, analyse de sources issues du renseignement, veille économique et concurrentielle, e-reputation, analyse de sentiments pour mieux comprendre les comportements des clients), le LDK de SYSTRAN vous permet de traiter et d'analyser des contenus multilingues structurés ou non, tels que les contenus générés par les utilisateurs, les contenus publiés sur les réseaux sociaux ou sur le Web.

Vous êtes maintenant parfaitement équipé pour créer des applications de veille et d'intelligence économique efficaces qui séduiront vos clients en recherche d'innovation !

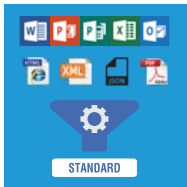
Ces services linguistiques sont distribués sous la forme d'un SDK (**Software Development Kit**). Nécessitant peu de ressources mémoire, les intégrateurs et développeurs externes peuvent ainsi aisément les embarquer dans leurs propres applications et services.

Prérequis

Librairies écrites en C++ nécessitant peu de ressources mémoire, utilisables par les langages de développement les plus courants (Java, .NET, Javascript, etc)

Systèmes d'exploitation supportés

Windows, Linux, Mac OS, iOS, Android



Filtrage de documents

Permet d'importer dans d'autres modules, des documents de différent format, et de recréer le document dans son format d'origine avec du contenu annoté ou modifié, sans changer la présentation.



Détection de la langue

Détecte automatiquement la langue dans laquelle le document est écrit grâce à l'analyse de mots ou de phrases spécifiques.



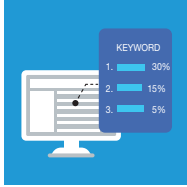
Segmentation et tokenisation

Segmente le texte en phrases, et les phrases en "tokens" (la plus petite unité de traitement).



Normalisation de la langue

Normalise du texte provenant de blogs, emails, chat ou de contenus générés par les utilisateurs et corrige aussi les fautes d'orthographe les plus courantes.



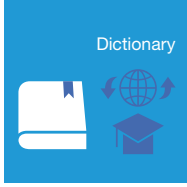
Catégorisation de documents

Détermine le sujet traité dans le document en se basant sur des modèles prédéfinis et affiche les mots clés les plus récurrents.



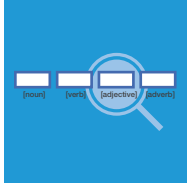
Reconnaissance d'entité nommée

Après analyse du contenu d'un document, reconnaît automatiquement et affiche les noms de personnes, les lieux, les nombres, les dates et les noms de société.



Dictionnaire

Permet d'effectuer une recherche dans un dictionnaire monolingue ou multilingue et d'avoir des informations contextuelles complémentaires telles que les fréquences d'usage, des contextes d'utilisation, des expressions et des exemples.



Analyse morphologique

Fournit une analyse morphologique pour chaque mot en proposant une liste des lemmes possibles et des composantes du langage pour une forme infléchie.



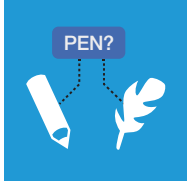
Analyse syntaxique

Fournit une analyse syntactique au niveau d'une phrase en intégrant des analyses linguistiques telles que l'identification de mot, identification des composantes du langage, analyses des composantes et des dépendances.



Translittération

Effectue la transcription de mots ou d'entités pour les langues possédant différentes formes d'écriture et détecte l'origine des noms propres.



Résolution des ambiguïtés

Choisit le sens d'un mot le plus approprié en fonction du contexte.