

Sayens - Offre de solution R&D

<https://www.sayens.fr/portefeuille-offres/offres-de-solutions-rd/technologie-des-langues/>

Technologie des Langues



Laboratoire d'adossement /Plateforme :

TIL (Centre Inter-Langues) , LIB (Laboratoire Informatique de Bourgogne)

Traduction automatisée

Contexte

L'industrie de la langue a pris un essor important autant en **traduction semi-automatisée** qu'en **compréhension écrite** (extraction automatique d'informations) comme **orale** de la langue (enceintes connectées etc.) ou en **production semi-automatisée de textes** (outils d'écriture prédictive, correction semi-automatisée, etc.) ou de **dialogues** (assistant conversationnel etc.).

Il est alors nécessaire de développer des **technologies informatiques**, des **modèles linguistiques** adaptés et des **corpus d'apprentissages** innovants et à la pointe pour rester compétitif sur ce marché.

Solution / Expertise

- Construction de corpus enrichis d'annotations sémantiques exploitables informatiquement :
 - **Collecte** de données
 - **Structuration** du jeu de données
 - Ajouts de tags sémantiques permettant des **analyses sur le sens plutôt que la forme**
- Extraction et quantification **des récurrences de sens ou de forme** d'un corpus de documents :
 - Analyses sommaires par un logiciel spécialisé
 - Analyses quantitatives approfondies par un linguiste
- Etude cognitive des récurrences pour en déduire une **architecture hiérarchisée** :
 - **Affectation** de structures socio-cognitives à ces récurrences
 - **Mise en relation** des structures cognitives
- **Entraînement d'un algorithme** (basé sur une intelligence artificielle) de traitement automatisé du langage :
 - **Génération** d'un algorithme
 - **Apprentissage** supervisé

Les projets peuvent être réalisés sur un vaste choix de langues.

Bénéfices

- Une méthodologie adaptée à vos besoins
- Une expertise européenne reconnue
- Des méthodes et des outils à la pointe de la connaissance internationale
- Un savoir-faire de plus de dix ans entre linguistes et informaticiens :
 - Thèse sur contrat doctoral i-site BFC autour de la **rédaction automatique de résumés**
 - Thèse sur contrat doctoral financement région BFC sur la modélisation de règles pour la **conversion de textes en français standard vers du français facile à lire et à comprendre**
 - Thèse sur contrat CIFRE sur l'amélioration de l'**idiomaticité d'un assistant conversationnel en plusieurs langues**

Contact : matthieu.bach@sayens.fr