

[\(https://unsplash.com/\)](https://unsplash.com/)

E. CABRIO

# Traitement automatique du texte en IA

En intelligence Artificielles (IA), le Traitement Automatique du Langage Naturel (TALN) est une discipline qui a pour objectif de modéliser, grâce à l'informatique, le langage qu'il soit écrit ou parlé. Les technologies TALN sont présentes, de manière grandissante, dans divers systèmes grands public (par ex. Google, IBM Watson, Facebook, Apple Siri).

 S1     3 ECTS     24h     OPT     [elena.cabrio@univ-cotedazur.fr](mailto:elena.cabrio@univ-cotedazur.fr) (<mailto:elena.cabrio@univ-cotedazur.fr>)

L'objectif de cette UE est de présenter les principales méthodes d'IA symbolique et d'apprentissage automatique permettant d'analyser, de générer, d'exploiter et de produire des documents en langue naturelle.

Le traitement automatique du langage naturel (TALN) est une discipline qui se trouve à la frontière de l'informatique et de la linguistique, et qui s'inscrit dans le domaine de l'intelligence artificielle. Elle regroupe l'ensemble des recherches et développements visant à modéliser et reproduire, à l'aide de machines, la capacité humaine à produire et à comprendre des énoncés linguistiques dans le but de communiquer. Il sera donc question ici de langage humain, d'où l'adjectif naturel, et non pas de langage formel. Pourquoi s'intéresser à l'automatisation du

traitement du langage naturel ? Comme pour la plupart des champs de connaissance ressortissant de l'IA, on peut identifier deux sources principales de motivation : d'une part la volonté de modéliser le langage, afin de tester des hypothèses sur les mécanismes de la communication humaine ; d'autre part le besoin de disposer d'applications capables de traiter efficacement les informations contenues dans les sources écrites ou sonores aujourd'hui disponibles sous forme électronique (pages HTML, documents hypermédias, médias sociaux, etc). À quoi servent des compétences en TAL ? Les métiers que suscitent les industries de la langue sont multiples. Des nombreuses possibilités s'offrent dans les entreprises spécialisées ou ayant des secteurs spécialisés dans le développement d'outils TAL (Google, Yahoo, IBM Watson, Microsoft, Facebook, Apple Siri, Amason Alexa, Lucene, Orange, France Telecom...) pour la conception et la maintenance de logiciels, pour les services qu'elles proposent ou pour leurs besoins propres. L'objectif de cette UE est de présenter les problématiques posées pour le traitement automatique du texte et les principales méthodes d'IA symbolique et d'apprentissage automatique permettant d'analyser, de générer, d'exploiter et de produire du langage naturel. Nous nous limiterons quasiment exclusivement au traitement du langage sous forme écrite. Cette UE aborde notamment :

- Fondements du TAL, c'est à dire les différents niveaux de traitement nécessaires pour parvenir à une compréhension complète d'un énoncé en langage naturel (analyse morphologique, analyse syntaxique, analyse sémantique, et analyse de la pragmatique, du discours et du dialogue). Du point de vue de l'ingénieur, ces niveaux correspondent aux modules qu'il faut développer et faire coopérer dans le cadre d'une application complète de traitement de la langue.
- Applications : recherche et extraction d'information, résumé automatique, fouille de texte, détection de sentiments, extraction de thématiques, systèmes de Question-Réponse, systèmes de dialogue, etc.

## Modalités de contrôle des connaissances



### À PROPOS

L'unité pédagogique informatique regroupe les enseignants d'informatique de la faculté des sciences (licence, master et parcours MIAGE) de l'Université Côte d'Azur. [Plus › \(/\)](#)

### SERVICES

[Contact \(/contact/\)](#)

[RSS \(/feed.xml\)](#)

[Atom \(/atom.xml\)](#)

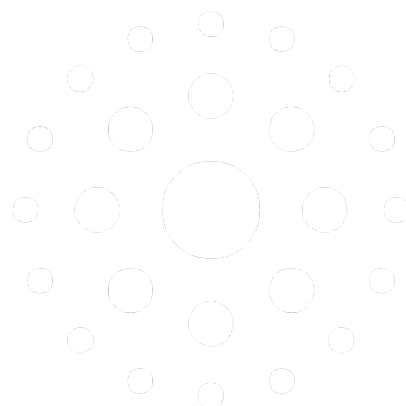
[sitemap.xml \(/sitemap.xml\)](#)



# SYSTÈMES NUMÉRIQUES POUR L'HUMAIN

## ÉCOLE UNIVERSITAIRE DE RECHERCHE

<https://ds4h.univ-cotedazur.fr>



UNIVERSITÉ  
CÔTE D'AZUR

<https://univ-cotedazur.fr>

Ce site est construit avec Jeekyll avec un thème graphique basé sur [FEELING RESPONSIVE \(HTTP://PHLOW.GITHUB.IO/FEELING-RESPONSIVE/\)](http://phlow.github.io/feeling-responsive/) par [PHLOW \(HTTP://PHLOW.DE/\)](http://phlow.de/).