



## L'analyse sémantique fine des textes et débats : introduction aux traitements automatiques et au modèle cartographique des Atlas sémantiques

### Objectifs et contenu

- Avoir un état de l'art sur les méthodes automatiques d'analyse sémantique de corpus
- Être capable de classifier les textes et d'en extraire les thèmes principaux
- Savoir analyser finement les contenus en mettant par exemple en évidence la variabilité sémantique des mots utilisés et des points de vue exprimés

Le traitement automatique des corpus de textes qui permet leur analyse sémantique est un domaine en plein essor. Les débats publics sous forme de textes libres (tels que ceux issus des plateformes mises en place dans la mouvance du Grand débat national) l'ont montré. Il s'agira de former les stagiaires aux méthodes utilisées en traitement automatique des langues.

Cette formation comprend un état de l'art des méthodes actuelles ainsi que des exercices pour leurs utilisations sur des exemples concrets. L'accent sera mis sur les modèles cartographiques favorisant une lecture synthétique des résultats et plus particulièrement sur celui des Atlas sémantiques..

- Etat de l'art sur les méthodes automatiques d'analyse sémantique de corpus et une introduction à leur interprétation
- Exercices adaptés aux thématiques d'intérêt des participants. Ces exercices seront choisis pour la mise en évidence des possibilités et des limites des modèles existants (TD)
- Présentation du modèle cartographique des Atlas sémantiques pour la représentation et l'analyse du sens
- Traitement exhaustif d'un cas. Il s'agira à partir d'un corpus de textes de tracer toutes les étapes de son traitement et d'interpréter les résultats produits (TD)

Les stagiaires pourront, s'ils le souhaitent, faire parvenir à la responsable de la formation, au moins quinze jours avant le début du stage, des types de corpus (domaines, genres, ..) qu'ils aimeraient analyser durant les TD à des fins pédagogiques sous réserve de l'accord de la responsable de la formation.

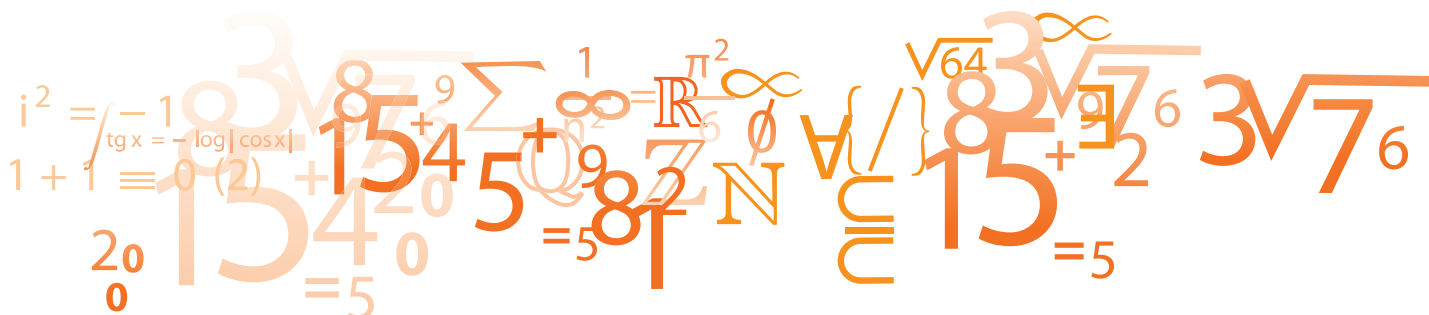
Alternance de cours (50 %) et de travaux dirigés (50 %)

### Intervenant(s)

Sabine Ploux (CAMS, EHESS)

Page personnelle : <http://cams.ehess.fr/sabine-ploux/>

Page du laboratoire : <http://cams.ehess.fr>





## Public visé

Toute personne intéressée par l'analyse automatique des corpus de textes et par l'interprétation des résultats de ces analyses

## Prérequis

Des notions de base en traitement automatique des langues ne sont pas nécessaires mais peuvent être utiles. Pas de prérequis en matière de programmation.

## Dates du stage, durée, emploi du temps et organisation

2 jours du lundi 30/03/2020 au mardi 31/03/2020.

De 4 à 10 stagiaires.

## Lieu

Paris.

## Coût pédagogiques

1000 Euros.

## Infrastructure ou équipement requis

Il est demandé aux stagiaires de venir avec leur propre ordinateur portable.

## A l'issue de la formation

Evaluation de la formation par les stagiaires.

Envoi d'une attestation de formation.

