



Annotation collaborative de corpus pour le TAL

présentation et organisation du cours

Karën Fort

karen.fort@sorbonne-universite.fr / <https://members.loria.fr/KFort/>



Présentation

Cours

Contrat d'enseignement

Plan du cours

Annotation collaborative de corpus pour le TAL

Descriptif

Avec l'essor de l'apprentissage supervisé, la formalisation linguistique, au sein du TAL, s'est (largement) déplacée des grammaires locales à l'annotation manuelle. Or, celle-ci reste un processus mal maîtrisé. Ce cours présente les enjeux de l'annotation manuelle de corpus pour le TAL, les problématiques liées à sa formalisation et à son évaluation et les solutions proposées, notamment la myriadisation (crowdsourcing), en particulier via des jeux ayant un but (GWAP).

Organisation du cours

- ▶ 8 séances de cours :
- ▶ évaluation : contrôle continu intégral
 - ▶ 1/2 : contrôle continu
 - ▶ 1/2 : examen final

Plan du cours

1. introduction et motivations
2. formats et outils
3. accords inter-annotateurs
4. solutions
5. *crowdsourcing* : typologie(s), légendes et réalités
6. Jeux ayant un but (*GWAP*) et gamification
7. dimensions de complexité de l'annotation