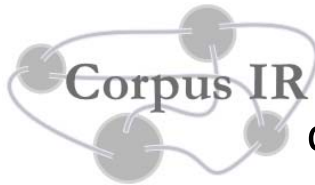


## Corpus-écrits GT7 "Nouvelles formes de communication –nouv-com- " : bilan 2012



- 21 membres, 14 unités de recherche (présentations sur Wiki)
- Communication, information
  - Une liste de discussions (<https://groupes.renater.fr/sympa/info/corpus-ecrits-nouvcom>),
  - un wiki (<https://groupes.renater.fr/wiki/corpus-ecrits-nouvcom/>)
- 2 réunions
  - en ligne, 31 mai 2012 (12 participants)
    - présentation sélection projets participants ayant conduit à production corpus (SMS, Twitter, blogues, forums, clavardage, corpus d'apprentissage, etc.)
  - en présentiel 28 juin (18 participants)
    - Tour de table : objectifs / thématiques de recherche des participants
    - Présentation exemple d'annotations morpho-syntaxiques sur corpus non standard
    - Travail groupe 1 : collecte et structuration des données en vue d'en faire des corpus échangeables et analysables manuellement et automatiquement ;
    - Travail groupe 2 : diffusion, partage, pérennisation des corpus nouv-com
    - Perspective 2013
  - Comptes rendus réunions et documents (diaporamas) sur Wiki



## Corpus-écrits GT7 "Nouvelles formes de communication –nouv-com- " : perspectives 2013



- Rassemblement, structuration commune, intégration et diffusion de corpus à partir d'une plate-forme unique (corpus déjà disponibles chez différents membres du GT7) comprenant nouvelles formes de communication (Twitt, blogues, textos, clavardage / chats, forums, etc.)
  - Diffusion en accès libre, respect standards (structures, métadonnées)
  - Intégration dans Corpus de Référence du Français
  - Recherche soutien logistique pour plate-forme au TGE-Adonis
  - Intégration métadonnées dans CLARIN
- Constitution d'un groupe de travail pour l'intégration des structures des différentes formes de nouvelles formes de communication dans la TEI
- Dépôt de projets ANR-Corpus entre plusieurs membres, plusieurs UR du GT7 pour pousser recherches, en particulier pour étiquetage, annotations
- Recension corpus ou données déjà existantes chez tous les membres du GT7 pour dépôt métadonnées dans CLARIN
- Formations organisées par Corpus-écrits sur thèmes tels que :
  - journée sur processus d'annotation : à partir de corpus disponibles à l'avance et d'outils (logiciels libres d'annotation), mettre en oeuvre une chaîne de traitement, avec segmentation, annotation morpho-syntaxique, voire plus
  - TEI