



Programme de la Journée d'étude

Bruit de fond ou valeur ajoutée ?

Gérer le bruit lors des traitements informatiques des corpus linguistiques

Accueil

8h00-8h30 Arrivée et accueil des participant.e.s. Un buffet de bienvenue est prévu.

8h30-8h45 Présentation de la journée d'étude

Session 1 - Axe n°1 : bruit et acquisition des données en dialectologie

8h45-9h30 Invitée : O. Nahli.

9h30-10h00 V. B. Lanza. *On orthography and normalization: the case of Judeo-Arabic.*

10h00-10h30 W. Remysen. *Les humanités numériques au service du patrimoine linguistique du Québec : l'exemple de la numérisation des fiches dialectologiques de Gaston Dulong.*

10h30-10h45 Pause café n°1

Session 2 - Axe n°2 : bruit et constitution des corpus d'apprenants écrits et oraux

10h45-11h30 Invité : L. Liégeois. *Des bruits dans mon corpus : des données à réduire au silence, à atténuer ou à écouter attentivement ?*

11h30-12h00 M. Barletta, C. Ponton. *La question de la normalisation des écrits scolaires pour leur traitement automatique. Le cas de l'omission de mots.*

12h00-12h30 S. Buchwalter et Q. Li. *Enjeux liés au traitement des erreurs dans la constitution de corpus écrits : l'exemple de corpus d'apprenants turcophones et sinophones de FLE*

12h30-14h00 Pause de midi (buffet)

Session 3 - Axe n°3 : bruit et annotation manuelle et automatique

14h00-14h30 J. Noblet. *Accord inter-juge et bruit dans l'annotation de la composante évaluative de l'émotion*

14h30-15h00 C. Koudoro-Parfait, G. Lejeune, L. Petrovic. *Correction automatique des interférences OCR dans la reconnaissance d'entités nommées spatiales : réel gain ou perte de l'information ?*

15h00-15h30 M. Maarouf. *Analyses fines de textes techniques bruités : stratégies de contournement.*

15h30-15h45 Pause café n°2

15h45-16h30 Invitée. Bénédicte Pincemin. *Gérer le "bruit" dans les corpus en textométrie : retour d'expérience et propositions*

16h30-17h30 Conclusion de Journée d'étude : bilan et perspectives

Informations générales

- Lieu : la journée d'étude aura lieu dans la Salle J. Cartier, situé dans la Maison des langues et des cultures, sur le Campus de l'UGA (1141, Av. Centrale, 38610, Gières (Tram B, arrêt Bibliothèques universitaires).
- Participation à distance :
- Durée des interventions : chaque intervention comprend 20 minutes d'exposé et dix minutes de questions/réponses, pour permettre au public d'approfondir certains aspects des communications
- Accueil et pauses : le comité d'organisation prendra en charge l'accueil, les pauses café, le buffet de midi et les déplacements des conférenciers invités.
- Contacts :