Science Participative pour les Données et Corpus Linguistiques (ParCoL)
Crowd-sourced linguistic data collection and augmentation

Atelier organisé dans le cadre de CORIA-TALN
avec le soutien du GdR LIFT-2
30 juin 2025

Salle 9-051, Campus Saint-Charles
Marseille, France

Programme

Matinée	
9h00-9h15	Ouverture des portes
9h15-9h30	Mot d'accueil des organisatrices
9h30-10h30	**Conférence invitée** *Sciences participatives 2.0 : Un changement profond dans la manière de produire des données*. Romain Julliard. Sorbonne U., MNHN, CNRS, CESCO, Paris, France.
10h30-11h00	☕ Pause-café
11h00-12h30	**Table ronde** Venez discuter de vos projets et de ceux des autres avec **Romain Julliard**, chercheur et directeur de la publication de SCIENCE ENSEMBLE (Sorbonne U., MNHN, CNRS, CESCO), **Antoine Srun**, président de Wikimédia France (Wikimédia France), **Lucas Lévêque**, contributeur Wikimédia et co-fondateur de Lingua Libre (Lug Numérique), **Florian Cuny**, contributeur Wikimédia (U. de Lorraine), **Marc Allassonnière-Tang**, chercheur (U. Paris Cité, MNHN, CNRS, EA), **Mathilde Hutin**, chercheuse (U. de Lorraine, CNRS, ATILF), **Alice Millour**, enseignante-chercheuse (U. Paris 8 Vincennes-Saint-Denis, LIASD), **Pierre Magistry**, enseignant-chercheur (ERTIM, INALCO), et **Ilaine Wang**, ingénieure de recherche (ERTIM, INALCO).
11h00-11h45	*Session 1. Méthodologies et logistiques de projets en cours* Présidente de session: Alice Millour, U. Paris 8 Vincennes-Saint-Denis, LIASD. - *Erreurs, idées reçues et plasticité en sciences participatives & Projet* Yezh-Ar-Vro *– La Langue du Pays.* Laure Turcati (Sorbonne U., Observatoire des Sciences de l'Univers Ecce Terra), Alice Millour (U. Paris 8, LIASD), Renaud Debailly (Sorbonne U., GEMASS), Karën Fort (U. de Lorraine, CNRS, Inria, LORIA), Asma Steinhausser (PatriNat, MNHN), Corentin Biets (CNRS-MNHN-UPVD, HNHP-PRETROP ; Eveha International) & Anne Dozières (Sorbonne U., MNHN, CNRS, CESCO). - *La science participative et l'ANR DiLSi*. Pierre Magistry (ERTIM, INALCO) & Ilaine Wang (ERTIM, INALCO).
11h45-12h30	*Session 2. Les données vocales participatives: Lingua Libre* Présidente de session: Mathilde Hutin, U. de Lorraine, CNRS, ATILF. - *Lingua Libre à l'ère de l'automatisation: l'I.A. au service du crowdsourcing d'un corpus oral*. Camille Lavigne (U. de Lorraine) & Florian Cuny (U. de Lorraine). - *LiLA : Outil d'augmentation automatisée des données vocales participatives de Lingua Libre*. Mathilde Hutin (U. de Lorraine, CNRS, ATILF), Marc Allassonnière-Tang (U. Paris Cité, MNHN, CNRS, EA), Lucas Prégaldiny (contributeur Wikimédia) & Lucas Lévêque (contributeur Wikimédia).

Après-midi	
12h30-14h00	🍴 Pause déjeuner
14h00-16h00	**Datathon** Formation express et contribution active à des outils et plateformes collaboratives.
14h00-15h00	*Wikidata* Animé par Lucas Lévêque (Lug Numérique). Présentation et versement de données... Venez verser votre bibliographie dans Wikidata, la base de connaissance qui alimente Wikipédia, le wiktionnaire, et tant d'autres !
15h00-16h00	*Wikisource et Dicothèque* Animé par Florian Cuny (U. de Lorraine) Venez découvrir Wikisource, la bibliothèque libre et collaborative, soutenue par la Wikimedia Foundation, et la Dicothèque, le comparateur de dictionnaires collaboratif !
14h00-16h00	*Lingua Libre* Animé par Mathilde Hutin (U. de Lorraine, CNRS, ATILF) En parallèle des deux datathons, vous aurez la possibilité de venir contribuer à Lingua Libre, la médiathèque participative de Wikimédia France.
16h00-16h30	☕ Pause-café
16h30	Fin

N'hésitez pas à vous joindre à nous uniquement le matin ou l'après-midi si vous le souhaitez !

Les inscriptions aux Ateliers CORIA-TALN sont ouvertes jusqu'au 23 juin !