

Journée Jeunes TALents 2024 - Programme

Jeudi 5 décembre 2024

13h00	13h15	Angèle Barbedette / Alain Couillault	Présentation des associations APIL et ATALA et de la Journée Jeunes TALents
Créneau		Intervenant(s)	Établissement public / Entreprise
13h15	13h45	<p>Céline Poudat, Christophe Parisse</p> <ul style="list-style-type: none"> - Deux stages autour du projet Open French Corpus: gestion des conversions / enrichissement des outils de recherche / travail sur les métadonnées et l'interface de l'OFC - Un stage: Projet Annotation - amélioration de TACTEO - développement et/ou conception de fonctionnalités / participation aux projet Inception - CORLI-GUM en cours - Un stage Citation de corpus - Un stage Bonnes pratiques d'analyse et de standardisation des corpus CMC 	CORLI
13h45	14h15	<p>Mathieu Constant</p> <ul style="list-style-type: none"> - Identification automatique de propriétés linguistiques d'expressions polylexicales - Annotation sémantique de définitions lexicographiques — application au Trésor de la Langue Française - Catégorisation onomasiologique du dictionnaire étymologique du français F.E.W. (Französisches etymologisches Wörterbuch) 	ATILF
14h15	14h30	<p>Mathilde Jouvel-Triollet</p> <ul style="list-style-type: none"> - Ingénierie de prompts d'un LLM dans le contexte de l'analyse contractuelle 	Viasema
14h30	15h	<p>Benoît Favre</p> <ul style="list-style-type: none"> - Exploring Complexity Factors in Multi-Hop Conversational QA https://pageperso.lis-lab.fr/frederic.bechet/talep/CALORDIAL_stage_LIS_2025.pdf - Multimodal conversation script generation, https://www.lis-lab.fr/wp-content/uploads/2024/11/Internship_-multimodal-conversation-script-generation-2.pdf - Child-Oriented AI: Advancing Effectiveness and Transparency in Interaction 	TALEP (LIS)
15h	15h30	<p>Farah Benamara</p> <ul style="list-style-type: none"> - LLMs jailbreaking (Create@CNRS Singapore) - LLMs-based continual learning and application in crisis management from social media (Create@CNRS Singapore) - LLMs for knowledge graph validation (IRIT-Toulouse) - Text classification in low-resource languages (IRIT-Toulouse) 	IRIT

15h30	16h	<p>Alex Cristia</p> <ul style="list-style-type: none"> - Extending whisper to perform voice type diarization - Voice type diarization frameworks and self-supervised learning models - Language development among Tsimane' children - Developing an ML-based toolkit for documenting language acquisition ecologically in under-resourced and multilingual settings - Data management for child speech corpora from under-resourced languages 	LAAC (LSCP)
PAUSE DE 5 MINUTES			
16h05	16h25	<p>Camille Guinaudeau</p> <ul style="list-style-type: none"> - Reference-Free Evaluation of Image Descriptions for Visually Impaired Users - Semi-Automatic Annotation of Conversations in Audio-Visual Documents 	LISN
16h25	16h40	<p>Philippe Suignard</p> <ul style="list-style-type: none"> - Production de petits modèles de langue (SLM) pour des domaines de spécialité 	EDF
16h40	17h	<p>Tom Bourgeade</p> <ul style="list-style-type: none"> - Adaptation de LLMs aux entretiens patient-docteur en français (https://recrutement.inria.fr/public/classic/fr/offres/2024-08399). - Verbalisation des Comportements Sociaux Audiovisuels pour le Dialogue Multimodal (https://recrutement.inria.fr/public/classic/fr/offres/2024-08393). 	LORIA
17h	17h15	<p>Hugues de Mazancourt</p> <ul style="list-style-type: none"> - Extraction d'informations pour un système d'intelligence stratégique territorial 	Datapolitics
17h15	17h30	<p>Maria Boritchev</p> <ul style="list-style-type: none"> - A computational socio-linguistic exploration of the Ethics of AI landscape 	Telecom Paris
17h30	17h45	<p>Bruno Palmino</p> <ul style="list-style-type: none"> - Mise en place de la brique open source Rasa dans le cadre du développement d'agents virtuels dans la santé 	Liaison Vocale
17h45	18h05	<p>Catherine Domingues</p> <ul style="list-style-type: none"> - Cahiers citoyens - Diagnostic et correction d'un corpus océréisé à l'aide de méthodes de TAL - Le thème de l'énergie dans la consultation citoyenne "le Grand Débat national" 	IGN
18h05	18h20	<p>Amalia Todirascu</p> <ul style="list-style-type: none"> - Annotation des expressions polylexicales dans des corpus d'apprenants 	LiLPa
18h20	18h30	<ul style="list-style-type: none"> - Arc Vector Representation for NLP tasks - Latent-Annotation Formalisms via VAE and Diffusion 	LIPN
18h30	18h35	Comité d'organisation	Conclusion