

NTeALan Research and Development

Conférence annuelle sur les technologies linguistiques et le TAL appliqués
aux langues peu dotées d’Afrique

NTeALan Group / Douala (Cameroun)

20 janvier 2021

1 Résumé

Thème	Avancées et défis du TAL pour les langues africaines : où en sommes-nous aujourd’hui ?
Date	Vendredi, 19 février 2021
Horaire	9h à 14h GMT
Invités	<ul style="list-style-type: none">— des chercheurs experts en TAL pour les langues africaines— des enseignants de langues africaines— des entrepreneurs
Lieu	Centre NTeALan sis au Parcours Vita (makepe Douala) et en ligne sur [https://webinar.nteabot.net/b/elv-tq4-oor-vrj], la page Facebook et Instagramm de NTeALan

2 Présentation de NTeALan

Créée en 2017 à Douala au Cameroun par Elvis Mboning et Jean-Marc Bassahak, l’association NTeALan (New Technologies for Africans Languages) œuvre pour l’émer-

gence numérique des langues africaines en construisant des communautés de contributeurs volontaires autour des langues et cultures africaines d’une part, et des technologies du traitement automatique des langues et de intelligence artificielle d’autre part. NTeALan est essentiellement formée de bénévoles passionnés de langues africaines et désireux de contribuer à leur développement. Toutefois, NTeALan travaille en partenariat avec des ministères (au Cameroun), des laboratoires de recherche, des entreprises et également des entrepreneurs représentant différents secteurs d’activités.

3 Description du workshop

Dans le cadre des activités de son assemblée générale marquant la clôture de l’année 2020 et le début de l’année 2021, l’association New Technologies for African Languages organise pour la seconde fois un workshop scientifique autour des problématiques du TAL et des langues peu dotées d’Afrique. Cette édition se déroulera le 19 Février 2021 simultanément en ligne sur la plateforme de webinar de NTeALan¹ et au siège de l’association (centre multimédia NTeALan), situé à Douala-Makepe, Tradex Parcours vita (en face de « Majestic Pressing»). Ce workshop regroupera la grande communauté des membres de l’association vivant au Cameroun et à l’étranger, des invités chercheurs, des partenaires et la communauté scientifique.

Ce workshop a pour but de réunir des professionnels, chercheurs et experts en langues africaines et en traitement automatique du langage naturel (TALN) dont les travaux de recherche sont axés sur les techniques du Machine Learning, la lexicographie électronique appliquée au TAL et la pédagogie/didactique des langues. Nous discuterons également des principaux défis qui se posent dans ce contexte pour la constitution des corpus de données en langues africaines, pour l’outillage des langues africaines. Nous définirons les orientations possibles pour les futurs progrès.

En résumé, ce workshop est l’occasion pour les intervenants et l’auditoire de :

- Passer en revue les travaux marquants sur la constitution des jeux de données utiles au Traitement Automatique des Langues Africaines (TALA).
- Présenter les dernières avancées, en matière de TAL, appliquées sur les langues africaines.
- Présenter les plateformes, les jeux de données, les logiciels actuellement disponibles sur le marché pour traiter les langues africaines.
- Évaluer et challenger les méthodes, techniques du TAL actuellement utilisées pour traiter les langues africaines.

1. <https://webinar.nteabot.net>

- Dynamiser les industries numériques propices à la digitalisation des langues africaines

Les intervenants pourront s’inspirer des thèmes suivants pour construire leurs sujets lors du workshop (liste non exhaustive) :

- Lexicographie électronique
- Codification et Xmlisation des dictionnaires
- Représentation des connaissances
- Méthodes de constitution des ressources lexicales
- Tokenisation, Lemmatisation, POS
- Application des données lexicographiques pour les outils TAL
- Traduction automatique des langues africaines
- Reconnaissance et synthèse vocale en langues africaines
- Classification sémantique de documents
- Synthèse vocale / Reconnaissance audio
- Identification des entités nommées présentes dans les langues africaines
- Système de question-réponse, Chatbot, NLU
- Extraction de données textuelles à partir des images, PDF, etc

4 Intervenants

Le workshop se tiendra le **vendredi 19 Février 2021** de **9h à 14h** en ligne sur la plateforme de webinar de l’association² et sera retransmise en direct au Centre NTeALan, sur la page Facebook et Instagram de NTeALan³. Les invités sont présentés dans le tableau ci-dessous (liste non-exhaustive).

5 Comité scientifique :

- Pr Jules Assoumou
- MCF Ndibnu Messina
- Dr Ornella Wandji
- M Elvis MBONING

6 Contact :

Contacts : levismboning@ntealan.org ; workshop@ntealan.org

2. <https://webinar.nteabot.net/b/elv-tq4-oor-vrj>

3. <https://www.facebook.com/NTeaLan.org>

Intervenants	Institutions	Thèmes	Sessions
Ouverture du workshop			9h00 - 9h15 UTC+1
Pr Assoumou Jules	Chef du Département de linguistique et langues africaines à l'Université de Douala	La question linguistique dans l'émergence de l'Afrique	9h15 - 9h45 UTC+1
Pr Sonja Bosch ⁽²⁾ & Dr. Gertrud Fass ⁽¹⁾	⁽²⁾ Professeure au Département de langues africaines de l'Université d'Afrique du Sud (UNISA)	An Learners' Dictionary for Zulu	9h45 - 10h15 UTC+1 <small>(10h45 - 11h15 UTC+2, heure de Johannesburg)</small>
MCF Ndibnu Messina	Chercheur à l'ENS et à l'Université de Yaoundé I	Le lexique des langues minoritaires africaines et la construction des corpus en TAL	10h15 - 10h45 UTC+1
Dr. Gertrud Fass ⁽¹⁾ & Pr Sonja Bosch ⁽²⁾	⁽¹⁾ Enseignante / Chercheur à l'Université de Hildesheim	Working towards a Zulu verb valence lexicon	10h45 - 11h15 UTC+1 <small>(10h45 - 11h15 UTC+1, heure de Berlin)</small>
Pause			11h15 - 11h30 UTC+1
M. Elvis Mboning	Président de l'association NTeALan	Les dictionnaires collaboratifs de NTeALan : quels avantages pour le traitement automatique des langues africaines ?	11h30 - 12h00 UTC+1
Mme Zouleiha Alhadji & Dr. Paul Dayang ⁽³⁾	Doctorante à l'Université de Ngaoundéré	A suffix-stripping Algorithm And transducers for the peul language	12h00 - 12h30 UTC+1
Dr. Paul Dayang ⁽³⁾ & M. Jules Paulin Bayang Souloukna ⁽⁴⁾	⁽³⁾ Chercheur sénior à l'Université de Ngaoundéré & ⁽⁴⁾ Doctorante à l'Université de Ngaoundéré	Computing perplexity values for under-resourced languages using ngram and deep learning approaches	12h30 - 13h00 UTC+1
Dr. Rodrigue Tchamna	Membre de l'association Resulam et Chercheur associé au City College de New York	Présentation des Travaux de Resulam, Résurrection des langues maternelles ancestrales, pour la revitalisation des langues africaines	13h00 - 13h30 UTC+1 <small>(7h00 - 7h30 UTC-5, heure de New York)</small>
Groupe de travail	Discussions / échanges thématiques		13h30 - 14h00 UTC+1
Conclusion et Fin			14h00 UTC+1