

Communiqué de presse



L'AVENIR EST FAIT D'AUDACE

Erganeo et Emvista signent une licence pour l'exploitation d'une ressource linguistique unique en son genre.

Erganeo a signé en mars 2022 un contrat de transfert de licence technologique avec la société Emvista, éditrice de logiciels, concernant une ressource linguistique : le French Tree Bank.

L'analyse et la reproduction du langage naturel nécessitent une compréhension et une structuration claire du sens de la phrase. Le corpus *French Tree Bank* (1996-2016) est une ressource lexicale et syntaxique unique en son genre, proposant 664 500 tokens (mots) et 21 500 phrases richement annotées (catégories, groupes et fonctions syntaxiques) par des linguistes. Elle répond notamment aux problématiques liées au Traitement Automatique du Langage Naturel (TAL), au web sémantique, à l'intégration du langage dans un moteur de recherche dialogue homme-machine (chatbots) ou encore à la vérification orthographique. Le corpus est utilisé dans plusieurs centaines de laboratoires de par le monde à des fins de recherche ou de traitement automatique du français.

Constitué à partir d'articles du quotidien *Le Monde*, le corpus *French Tree Bank* (corpus arboré pour le français) a été déposé par Université Paris Cité, qui a confié à Erganeo la valorisation de cette ressource linguistique. La technologie, développée par les chercheurs Anne Abeillé (professeure, UMR 7110 CNRS Université Paris Cité) et Lionel Clément (doctorant et maître de conférences à l'Université de Bordeaux), a donné lieu à un dépôt de brevet en 2008.

Le *French Tree Bank* est aujourd'hui transféré auprès d'Emvista, une société montpelliéraine éditrice de logiciels, qui développe une technologie à l'état de l'art pour analyser tout contenu textuel et en structurer l'information qui y est véhiculée. Cette technologie de structuration de l'information s'appuie sur de nombreux indicateurs linguistiques dont ceux annotés dans le corpus *French Tree Bank*. Ces indicateurs seront ensuite utilisés au sein de systèmes d'apprentissage et à base de connaissances. Le *French Tree Bank* est également une ressource de référence dans le milieu académique et est donc indispensable à l'équipe de R&D de Emvista dès lors que celle-ci souhaite comparer les performances des systèmes expérimentés.

Actuellement, le corpus est distribué gratuitement à des fins de recherche, et une version d'essai est disponible pour tous sur le site *French Tree Bank*.

[\(http://ftb.linguist.univ-paris-diderot.fr/\)](http://ftb.linguist.univ-paris-diderot.fr/)

À propos d'Erganeo – www.erganeo.com

Erganeo est une société d'investissement française spécialisée dans les innovations de rupture (DeepTech) à fort impact sociétal. Nous investissons au plus tôt pour sécuriser les nouvelles inventions des chercheurs, en amont de leur transfert à un industriel ou de la création de start-up, et ce dans de nombreux domaines scientifiques : Biotech, Infotech (Telecom, objets connectés, big data, IA), Eneritech (énergies nouvelles, chimie, matériaux) etc. Notre mission est d'accélérer et de simplifier les associations entre la Recherche et l'Industrie en faveur d'un progrès sociétal. Pour ce faire, nous finançons et nous accompagnons vers la réussite et la reconnaissance internationale la nouvelle génération de chercheurs-entrepreneurs français.

Membre du Réseau SATT, Erganeo puise les bases d'un futur souhaitable dans l'écosystème francilien, un vivier d'innovations riche de 20 000 chercheurs répartis dans plus de 370 laboratoires de recherche de pointe. Depuis sa création, Erganeo a investi plus de 42 M € et ainsi contribué à signer 97 licences avec des entreprises de toutes tailles et à créer 31 start-ups.

Contact presse : Caroline Pontifice
Responsable de la communication et du marketing
caroline.pontifice@erganeo.com

À propos d'Emvista - www.emvista.com

Emvista est un éditeur de logiciels. Créée à Montpellier en 2018 et forte de son équipe de R&D en Traitement Automatique du Langage Naturel, la société propose des produits innovants qui s'appuient sur des technologies à l'état de l'art. Son produit phare, l'assistant de gestion d'e-mails, permet à ses utilisateurs d'être plus efficaces dans la gestion de leurs e-mails en repérant dans la masse d'information reçue toutes les informations pertinentes (par exemple les demandes d'information et d'actions qui nécessitent une action en réponse). Emvista commercialise de nombreux autres services d'analyse de textes tels que l'analyse d'opinions et d'émotions, l'anonymisation de textes sensibles ou encore l'extraction d'entités.

Contact presse : Cédric Lopez - Directeur de recherche
cedric.lopez@emvista.com