

DOSSIÈR DE PREMSA



Plateforme des technologies de la langue pour l'occitan

La Farga numerica est une plateforme collaborative d'accès, de valorisation, d'échange et de développement des ressources et outils TAL (traitement automatique des langues) et informatiques pour la langue occitane. Elle répond à un besoin identifié de la communauté et des développeurs en termes d'accès à des données qualifiées, de mutualisation de la veille, de coordination des travaux et de partage des développements.

Inspirée des **espaces de mutualisation et de collaboration** de la communauté du logiciel libre (distributions *Linux*, suites logicielles, etc.), *La Farga numerica* se veut être à la convergence des initiatives. Grâce à son interface évolutive, elle offre **un bouquet de ressources et de fonctionnalités en direction des développeurs et des contributeurs** : inventaire qualifié et actualisé des ressources, feuille de route numérique, standards informatiques adaptés à l'occitan, contact des différents acteurs et projets structurants, appels à projets, appels à contribution, messagerie instantanée, publications scientifiques, veille technologique.



Accéder à *La Farga numerica* : lafarganumerica.eu

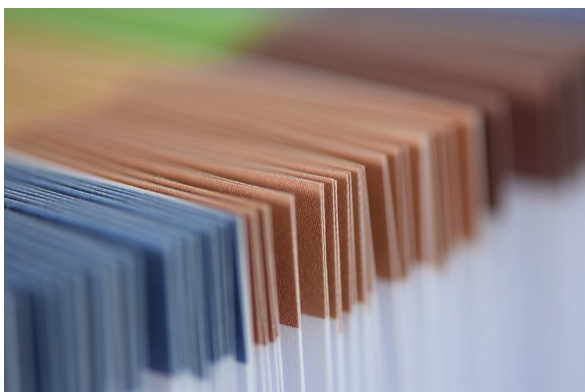
Dossier de presse numérique à l'adresse premsa.locongres.com/lafarga.

Avec le soutien de

/ Un espace pour les développeurs

L'objectif de *La Farga numerica* est de **faciliter le travail** de ceux qui développent des outils numériques en occitan, **en recensant** en un seul endroit les ressources à leur disposition. Il peut s'agir de données sous licence ouverte, de logiciels libres, de modèles pré-entraînés, d'API qu'on peut utiliser pour programmer... mais aussi de vocabulaire technique ou de standards qui permettent de **faire de l'occitan une langue à la fois sujet et objet du numérique**.

Inventaire des ressources et outils



Des corpus, des logiciels, des services... pour ceux qui travaillent dans le TAL ou l'informatique occitane, avec un sceau de qualité et également une section spécifique pour les API et une pour les nouveautés.

Standards informatiques



Des propositions d'étiquettes, de formats... standardisés et interopérables avec les normes internationales, et également adaptés aux spécificités de la langue et la culture occitanes.

Vocabulaire technique en occitan



Le vocabulaire de l'informatique et des nouvelles technologies, pour tous ceux qui produisent du contenu en occitan à propos des technologies de la langue, comme ceux qui traduisent des logiciels.

Tutoriels sur le numérique occitan



Des petits tutoriels, pour le grand public comme pour les initiés, pour mieux intégrer l'occitan dans leurs pratiques numériques : comment taper les accents, passer Firefox en occitan, installer un correcteur...

/ Le lieu de rencontre de la communauté

Quand certaines langues ont des multinationales qui œuvrent à faire avancer leur traitement automatique, d'autres ne disposent pas de milliers de personnes qui travaillent à leur informatisation. L'occitan, comme de nombreuses autres langues peu dotées, **doit une grande part de ses avancées au contributif, au collaboratif et à l'open source.**

La communauté qui s'occupe de TAL occitan est éparpillée dans plusieurs régions d'Occitanie, et même pour certains à l'international. *La Farga numerica* veut faire le lien entre tous ces acteurs qui travaillent, parfois, chacun dans son coin. Elle veut être leur lieu de rencontre, d'échanges pour favoriser l'émulation et le bouillonnement d'idées.

Pour ce faire, elle propose un espace pour tous ceux qui ont envie de participer à la grande aventure des technologies de la langue pour l'occitan. Ils y trouveront des notices qui donnent des clés pour contribuer dans plusieurs domaines, des appels à participations, des contacts pour aider ceux qui ont un projet numérique...

Le Discord de *La Farga numerica*



La Farga numerica anime un serveur *Discord* sur lequel la communauté informatique occitane peut se retrouver pour échanger et travailler ensemble.

/ Le paysage du TAL occitan

Cela fait plusieurs années que des acteurs institutionnels et des contributeurs travaillent sur le traitement automatique de la langue occitane. Depuis la première feuille de route pour le développement numérique de l'occitan en 2014, ils essaient d'**œuvrer de façon coordonnée vers des objectifs communs.**

Il existe aujourd'hui une véritable dynamique dans ce domaine, avec des grands acteurs et des projets structurants, mais aussi une galaxie de contributeurs qui permettent à l'occitan d'avoir des ressources et des outils de qualité, dans un champ où plusieurs langues « peu dotées » ne sont pas aussi riches.

État des lieux du domaine



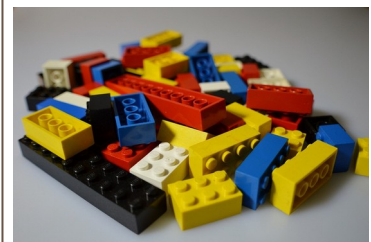
La situation actuelle pour les ressources et les outils du TAL occitan : traduction automatique, synthèse vocale, lexiques bilingues, corpus audios...

Les acteurs institutionnels



Quelles sont les structures derrière les outils que vous utilisez ? Qui œuvre à la constitution des ressources qui entraînent les IA ou qui nourrissent les algorithmes ?

Les principaux projets



Les acteurs se regroupent autour de plusieurs initiatives, pour y mutualiser leurs ressources et leurs compétences, et développer des services pour le grand public.

Les personnes ressources



Découvrez les « petites mains » des technologies de la langue en occitan, des professionnels et des bénévoles qui construisent les outils informatiques.

Publications scientifiques



Bibliographie de publications scientifiques qui concernent le TAL et la langue occitane.

/ Le Congrès

La Farga numerica est un portail édité par **Le Congrès**.

Le Congrès permanent de la langue occitane est l'organisme interrégional de régulation de l'occitan. Il œuvre dans les domaines de la linguistique et du TAL (traitement automatique de la langue).

Il produit des outils linguistiques numériques de référence (dictionnaires, conjugueurs, correcteurs orthographiques...), des applications pour le TAL (synthèse vocale, traduction automatique...) et des applications pour les mobiles (claviers prédictifs...).

Il a également des missions de régulation linguistique et de recherche scientifique appliquée.

Il est l'éditeur de *Dicodòc* (multidictionnaire occitan), *Revirada* (traducteur automatique occitan) e *Votz* (synthèse vocale occitane).



/ Linguatec IA

La Farga numerica est issue du projet **Linguatéc IA**.

Le projet « Linguatéc-IA, développement des connaissances en intelligence artificielle pour la numérisation de l'aragonais, du catalan, du basque et de l'occitan » a pour objectif de rechercher et développer de nouveaux modèles neuronaux applicables aux langues faibles en ressources, et construire une infrastructure linguistique transfrontalière intelligente facilitant la communication entre locuteurs de langues différentes et l'accès multilingue à l'information.

Linguatéc IA est cofinancé par :



/ Pour plus d'informations

Lo Congrès permanent de la lenga occitana
Ciutat - Creem
5-7 rue de la Fontaine
64000 Pau

premsa.locongres.com/lafarga

info@locongres.org

06 45 57 01 07