



LINGUA LIBRE



Enregistrer toutes les langues du monde, commune par commune...



L'équipée

Pierre-Antoine Le Page
Noé Gasparini
Olivier Quint
Aure Segulier
Rémy Gerbet
Hugo Lopez
Lucas Lévêque
Nicolas Vion
Mathieu Stumph Guntz
Manuel Fauveau
Félix Mitjans
Olivier Delaune
Antoine Lamielle





Il était une fois....

Un constat édifiant sur le wiktionnaire :

- Une sous-représentation grave des langues **régionales** de la France malgré des sources de plus **en plus** numérisées et disponibles.
- Un manque flagrant de données ouvertes pour **les** illustrer.
- L'inaccessibilité apparente des données de recherches sur le sujet.





Il était une fois...

Une discussion entre deux conférences avec Pierre-Antoine Le Page :

- Les langues régionales font parties des richesses culturelles les plus en danger en France. Leur disparition est actée pour beaucoup...
- Mais tout n'est pas perdu. Une veille précise et méthodique sur le net montre une vitalité insoupçonnée à certains endroits de la France.
- Nous pouvons faire quelque chose qui n'a jamais été fait précédemment grâce aux projets de la Wikimedia Fondation.



Il était une fois...

Le projet Shtooka :

- Un logiciel d'enregistrement massif des langues.
- Plus de 100 000 enregistrements venant d'une quinzaine de langues.
- Beaucoup ont déjà été téléversés sur Commons et illustrent les projets de la fondation (Wikipédia, Wiktionnaire...)
- Grâce à une technologie innovante, il permet d'enregistrer 1000 mots en moins d'une heure.



Il était une fois....

Étape 1 : Lancement du projet « Langues de **France** ».

Étape 2 : Prise de contact avec la DGLFLF.

Étape 3 : En coopération avec la DGLFLF, lancement **de** la réflexion sur un outil d'enregistrement des langues régionales. Contact avec Nicolas Vion, développeur de Shtooka.

Étape 4 : Novembre 2015 – Démarrage du projet « **Lingua Libre** ».



Parce qu'il faut prendre **le temps** :

- 3 ans de développement.
- Des contacts réguliers avec des linguistes **pour** calibrer le développement.
- Un développement effectué par des bénévoles volontaires et deux développeurs momentanément à plein temps.



Lingua Libre !

C'est un outil fonctionnel et performant :

- Enregistrement par liste de mots : Permet d'enregistrer plus de 1000 mots en 45 minutes avec les méta-données de base.

C'est actuellement plus de 30 000 enregistrements dans plus de 20 langues !



Lingua Libre !

C'est un gestionnaire complet de locuteurs et de locutrices pour un grand nombre de situations d'enregistrement. Avec suffisamment de métadonnées pour que chaque enregistrement soit de qualité linguistique, c'est-à-dire réutilisable pour la recherche.

Toujours en pensant à garder simple l'enregistrement pour qu'il soit accessible à tous.



Et le mieux, c'est un **exemple**:

Juste après cette présentation, une **démonstration en direct** des deux voies d'enregistrement.



Une aide internationale !

- 4 développeur.euses volontaires rencontrés à un hackathon (rassemblement de développeur) en Autriche sont intéressés.e.s pour internationaliser et développer l'outil.
- 2 Indiens, 1 Israélienne et 1 Allemand.
- Un apport de savoir-faire et de bonne volonté non-négligeable, mais une nécessité d'organisation de la part de l'équipe.



Remerciements !



Gau un Griis



L'Alsace en Version Originale

