

DiLAF

Université de Nantes

Numéro du projet: 2015-670

Demande de soutien

Information sur la structure porteuse du projet

Nom de la structure

Université de Nantes

Statut juridique de la structure

établissement public à caractère scientifique, culturel et professionnel

Date d'immatriculation

29/12/1961

Mission

Les missions de l'Université de Nantes sont définies dans le Code de la Recherche.

Article L321-1

La mission de ces établissements est de mettre en oeuvre les objectifs définis à l'article L. 112-1.

Article L112-1

La recherche publique a pour objectifs :

- a) Le développement et le progrès de la recherche dans tous les domaines de la connaissance ;
- b) La valorisation des résultats de la recherche au service de la société, qui s'appuie sur l'innovation et le transfert de technologie ;
- c) Le partage et la diffusion des connaissances scientifiques en donnant priorité aux formats libres d'accès ;
- c bis) Le développement d'une capacité d'expertise et d'appui aux associations et fondations, reconnues d'utilité publique, et aux politiques publiques menées pour répondre aux défis sociétaux, aux besoins sociaux, économiques et du développement durable ;
- d) La formation à la recherche et par la recherche ;
- e) L'organisation de l'accès libre aux données scientifiques.

Les établissements publics de recherche et les établissements d'enseignement supérieur favorisent le développement des travaux de coopération avec les associations et fondations, reconnues d'utilité publique. Ils participent à la promotion de la recherche participative et au développement des capacités d'innovation technologique et sociale de la Nation. Ces coopérations s'exercent dans le respect de l'indépendance des chercheurs et, en l'absence de clauses contraires, dans un but non lucratif. Les travaux de recherche menés dans le cadre de ces coopérations sont, en l'absence de clauses contraires, rendus publics et accessibles.

Zone(s) géographique(s) d'intervention

Europe de l'ouest

Afrique et Moyen Orient

Public(s) cible(s)

Les publics cibles de ce projet sont

- les élèves et enseignants dans les établissements d'enseignement

- les populations locutrices des langues des dictionnaires,
- les personnes apprenant les langues des dictionnaires,
- les chercheurs en traitement automatique des langues naturelles.

Nom du représentant légal

Olivier Laboux

Fonction dans la structure

Président de l'Université

Nombre de salariés

4260

Nombre de bénévoles actifs

0

Adresse du siège social**Rue et numéro**

1, quai de Tourville

Code postal

44000

Ville

Nantes

Pays

France

Téléphone fixe

02 40 99 83 83

Téléphone portable

02 40 99 83 83

Fax

02 40 99 83 00

Sigle/acronyme de la structure

Université de Nantes

Adresse e-mail

accueil.presidence@univ-nantes.fr

Site internet

<http://www.univ-nantes.fr>

Responsable du projet au sein de la structure**Civilité**

Madame

Nom

Enguehard

Prénom

Chantal

Fonction dans la structure

Maître de conférences en informatique

Adresse**Rue et numéro**

2, rue de la Houssinière

Complément d'adresse

BP 92208

Code postal

44322

Ville

Nantes Cedex 03

Pays

France

Téléphone fixe

02 51 12 58 55

Téléphone portable

06 25 17 59 19

Fax

02 51 12 58 12

Adresse e-mail

chantal.enguehard@univ-nantes.fr

Avez-vous déjà soumis un projet à la Fondation ?

Non

Comment avez-vous connu la Fondation ?

Internet

Description du Projet

Nature du projet

Science

Nom du projet

DiLAF

Contexte qui sous-tend la cause

Certaines langues sont dites "peu dotées" : les ressources linguistiques les concernant sont peu diffusées ou inaccessibles. C'est typiquement le cas dans de nombreux pays d'Afrique où

? la langue officielle (le plus souvent, le français ou l'anglais) n'est comprise que par une minorité,

? il y a plusieurs langues nationales véhiculaires largement utilisées mais peu dotées,

? le taux d'analphabétisme est important, jusqu'à 80% dans la zone sahélienne.

Des écoles bilingues/multilingues où sont utilisées la langue officielle et une ou plusieurs langues nationales se développent et de bons résultats sont observés par l'Académie Africaine des Langues (ACALAN) ou l'Organisation Internationale de la Francophonie. Au sujet de l'enseignement, le Plan Décennal du Développement de l'Éducation (PDDE) au Niger note également que "Les filles auraient de meilleurs résultats si elles étaient scolarisées dans des idiomes qu'elles maîtrisent, les langues nationales."

Description du projet

DiLAF est un projet de recherche visant la conversion de dictionnaires éditoriaux en des dictionnaires électroniques Les dictionnaires peuvent être consultés gratuitement en ligne ou hors ligne.

En effet, de nombreux dictionnaires bilingues de bonne qualité existent. Ils sont commercialisés sous forme éditoriale (papier). Du fait de la faiblesse économique de la population, ils sont peu vendus et ne remplissent donc pas l'objectif d'être des ouvrages de référence.

Les collaborateurs africains travaillant sur ce projet sont des professionnels : chercheurs en informatique ou en linguistique, lexicologues. Ils appartiennent à des institutions spécialisées dans les langues nationales. Il s'agit de l'INDRAP (Institut National de Documentation de Recherche et d'Animation pédagogiques) et de la DRELN (Direction de la Recherche et de l'Équipement des Langues Nationales) au Niger, et du CLAD (Centre de Linguistique Appliquée de Dakar), au Sénégal.

Les dictionnaires sont mis en ligne avec l'accord des auteurs. Ils sont alors protégés par une licence Creative Commons CC BY-NC-

SA 2.0 . Celle-ci autorise la consultation et la réutilisation non commerciale des dictionnaires.

Le projet poursuit plusieurs objectifs à long terme :

- ? lutter contre l'analphabétisme,
- ? réduire le taux d'échec scolaire,
- ? faire bénéficier les langues nationales des outils de Traitement Automatique des Langues Naturelles (TALN).

À moyen terme :

- ? faciliter l'accès aux dictionnaires, favoriser l'expression d'écrits en langue nationale, principalement chez les jeunes,
- ? encourager la production de pages web bilingues et dans les langues nationales,

À court terme, le projet financera l'ajout d'un dictionnaire (fulfulde), l'enrichissement de cinq dictionnaires (dont quatre sont déjà en ligne) : haoussa, kanouri, tamajaq, wolof et zarma. Ces langues sont parlées par plusieurs dizaines de millions de personnes.

Nombre de personnes bénéficiaires de ce projet

8

Lieu du projet

Ville et pays

Nantes, France

Quelle est l'originalité de votre projet, ses forces et faiblesses et ses facteurs de succès ?

FORCES

Le projet existe déjà et a donc montré sa faisabilité. Cinq dictionnaires sont en ligne (bambara, haoussa, kanouri, tamajaq et zarma), leur consultation est gratuite.

À notre connaissance il n'existe aucun site web offrant l'accès gratuit à des dictionnaires de langues africaines.

Le site web est largement consulté. Les dictionnaires ont été téléchargés plusieurs centaines de fois. Ce succès s'explique par l'absence de ressources linguistiques en ligne et de qualité concernant ces langues.

FAIBLESSES

Le projet, débuté au Niger et au Burkina Faso, ne peut continuer ainsi du fait de l'insécurité. Il faut donc déplacer les collaborateurs africains en Europe.

Deux missions par dictionnaire seront nécessaires. En effet, les collaborateurs africains doivent se libérer de leurs tâches habituelles pour les missions de travail à Nantes. Or, une fois en France, ils sont coupés de leurs réseaux locaux de travail qu'il serait parfois utile de consulter.

Suivi et évaluation

Chantal Enguehard, enseignante-chercheuse en informatique et spécialiste en TALN, assure la coordination et la direction scientifique du projet depuis ses débuts, en collaboration avec Mathieu Mangeot (enseignant-chercheur à l'Université de Grenoble).

Le projet verra l'accueil des collaborateurs africains (linguistes, lexicologues, formateurs en langue nationale) à Nantes afin de corriger et d'enrichir quatre dictionnaires déjà en ligne (haoussa, kanouri, tamajaq, zarma) et d'ajouter deux dictionnaires : fulfulde et wolof.

La durée du projet (sur 18 mois) permet d'organiser cet aller-retour. En effet, les collaborateurs sont des professionnels qualifiés, parfois universitaires, qui devront se libérer pour réaliser ces missions.

Le succès du projet pourra être mesuré par

- le nombre de notices enrichies ou révisées dans chacun des dictionnaires déjà en ligne,
- le nombre d'exemples traduits en français,
- le nombre de définitions traduites en français,
- le nombre de notices du dictionnaire fulfulde qui sera mis en ligne,
- le nombre de notices du dictionnaire wolof qui sera mis en ligne,
- le nombre de consultations en ligne des dictionnaires,
- le nombre de téléchargements des dictionnaires,
- les résultats d'un questionnaire de satisfaction (en ligne) visant les personnes ayant téléchargé un ou plusieurs dictionnaires.

Par ailleurs le projet a déjà donné lieu à plusieurs publications scientifiques. Voici deux références

- Mangeot, M. Enguehard, C. Des dictionnaires éditoriaux aux représentations XML standardisées. Chapter 8. in Gala, Núria et Michael Zock (dir.), Ressources Lexicales: Contenu, construction, utilisation, évaluation. xii, 364 pp. (pp. 255-290). 2013.
- Enguehard, C. Mangeot, M. LMF for a selection of African Languages. Chapter 7. in Gil Francopoulo (dir.), LMF: Lexical Markup Framework, theory and practice. Ed., Hermès science. Paris, France. 17 p. 2013.

Pérennité du projet

La direction et la pérennité du projet seront pris en charge par Chantal Enguehard, enseignante-chercheuse ayant un poste à l'Université de Nantes, et qui conservera ses activités de recherche sur les langues africaines.

Le domaine de recherche du Traitement Automatique des Langues Naturelles appliqué aux langues peu dotées est appelé à prospérer avec l'informatisation des langues du monde entier.

Le nom de domaine du site web www.dilaf.org a été acheté, en 2011 pour dix ans et sera prolongé le moment venu. Les institutions

africaines hébergeant les chercheurs impliqués dans ce projet sont chargés de mettre en œuvre la politique linguistique des États et ont donc vocation à poursuivre leurs activités.

Les ressources produites sont de simples fichiers informatiques ne nécessitant aucune maintenance. Elles sont donc pérennes. La licence Creative Commons permet la réutilisation gratuite de ces ressources (tant que le producteur initial est cité). Ces ressources ont vocation à se diffuser.

Organisation du projet

Le projet sera géré par l'Université de Nantes via le laboratoire d'Informatique de Nantes-Atlantique (LINA). Ce dernier se chargera d'organiser les missions des collaborateurs africains en son sein (accès à un bureau, à un ordinateur, aux ressources bibliographiques, etc.).

Pour chaque dictionnaire, deux missions d'un mois chacune sont nécessaires : une première mission permettra d'avancer le travail et de collecter les questions linguistiques ou lexicographiques qui ne peuvent être tranchées sur place. Les collaborateurs seront chargés de consulter des collègues afin d'apporter des réponses à ces questions.

Une seconde mission permettra de prendre en compte les réponses aux questions, d'ajouter éventuellement de nouvelles entrées aux dictionnaires et d'achever la révision.

Mathieu Mangeot (spécialiste des dictionnaires électroniques) co-animera les ateliers de travail avec Chantal Enguehard. Ils initieront des travaux de TALN sur les langues des dictionnaires.

Calendrier détaillé du projet

Date de lancement du projet

01/01/2016

Date de fin du projet

31/07/2017

État d'avancement du projet à la date de cette demande de soutien

Cinq dictionnaires sont déjà en ligne sur le site web DiLAF. www.dilaf.org

Le dictionnaire bambara-français est achevé. Les dictionnaires kanouri-français, haoussa-français, zarma-français et tamajaq-français devraient être révisés, complétés et enrichis. La traduction des exemples et des définitions en français permettra de constituer des corpus bilingues alignés. De tels corpus sont utilisés en traduction automatique. Ils seront également utiles aux futures recherches en TALN sur ces langues.

Un dictionnaire fulfulde-français et un dictionnaire wolof-français/anglais sont en cours de conversion pour être mis en ligne. Ces actions sont pour l'instant suspendues car la contribution d'un lexicologue ou d'un linguiste est indispensable.

Les dictionnaires téléchargés au format XML ne sont, pour l'instant, pas lisibles par de simples utilisateurs. Une amélioration de la consultation hors ligne est en cours de développement et sera mise en place en septembre 2015.

Autres commentaires

Le projet DiLAF a vocation à abriter d'autres dictionnaires. Il reste donc susceptible de connaître des développements ultérieurs sur d'autres langues.

Budget

Budget total du projet

121464 EUR

Montant du soutien financier demandé à la Fondation

75882 EUR

Affectation prévue de ce montant

- Missions des partenaires à Nantes (ateliers de travail de transformation des dictionnaires)
- Deux ordinateurs portables
- Clefs USB
- Batterie DELL

Combien d'autres partenaires sont impliqués ?

5

Partenaire 1

Nom

Université de Nantes

Montant accordé

15664 EUR

Partenaire 2

Nom

Université de Grenoble

Montant accordé

2457 EUR

Partenaire 3

Nom

INDRAP - Niger

Montant accordé

38527 EUR

Partenaire 4

Nom

CLAD - Sénégal

Montant accordé

9617 EUR

Partenaire 5

Nom

DRELN - Niger

Montant accordé

9617 EUR

Avez-vous d'autres besoins en dehors d'un soutien financier ?

Non

Téléchargement des documents administratifs à joindre au dossier

- [STATUTS version modifiée par le CA du 30 01 2015.pdf](#)
- [JO196112503.PDF](#)
- [LEGITEXT000006071190.pdf](#)
- [Rapport-Annuel-Financier_2014-complet.pdf](#)
- [budget.pdf](#)

Documents obligatoires

Statuts

Déclaration au Journal Officiel

Dernier rapport d'activité

Dernier rapport financier

Budget détaillé du projet

