

INSTRUCTIONS

Veillez faire remplir et soumettre le contrat utilisateur final en deux exemplaires originaux signés par la personne dûment habilitée à engager la responsabilité de votre organisation.

Veillez parapher chaque page du contrat (le signataire doit inscrire ses initiales sur chaque page).

L'utilisateur doit être une institution légale ou un département ou une section d'une institution et non un individu ni un projet. Le signataire de ce contrat doit être une personne habilitée par son institution à signer ce type de contrat.

En Annexe B, veuillez indiquer le lieu où les données seront exploitées.

En Annexe C, veuillez indiquer les tâches que vous réaliserez. L'accès à ces données est autorisé uniquement dans le cadre de ces tâches. Si, pour des raisons indépendantes de votre volonté, vous ne pouvez achever l'ensemble de ces tâches, veuillez nous en informer. Si, vous décidez ultérieurement d'effectuer une tâche supplémentaire, veuillez nous en informer et nous vous enverrons un addendum au contrat si des données complémentaires sont requises.

Les deux exemplaires originaux du contrat doivent être renvoyés par courrier à l'adresse ci-dessous :

*Valérie MAPELLI
ELDA / ELRA
55-57, rue Brillat Savarin
75013 Paris, France.*

*Tel : 01.43.13.33.33
Fax : 01.43.13.33.30*



55/57, rue Brillat Savarin
75 013 Paris
Tél. : 01 43 13 33 33
Fax : 01 43 13 33 30
Web : www.elda.org
Email : info@elda.org

CONTRAT UTILISATEUR FINAL DE RESSOURCES LINGUISTIQUES DANS LE CADRE D'UN PROJET D'ÉVALUATION

(Contrat n° LC/ELDA/ENDUSER/ETAPE/2011/.....)

Entre

ELDA (Agence pour l'Évaluation et la Distribution des Ressources Linguistiques), S.A. au capital de 250 000 euros, dont le siège social est situé : 55-57 rue Brillat Savarin - 75013 Paris, France, immatriculée au Registre du Commerce et des Sociétés de Paris : RCS Paris B 402 781 876 (95b147 95) représentée par Monsieur Khalid CHOUKRI agissant en sa qualité de PDG,

et

".....", dont le siège social est situé :.....

Ci-dessous dénommée "L'UTILISATEUR" et représentée par
.....,
agissant en sa qualité de
.....

Préambule :

La campagne ETAPE (Évaluations en Traitement Automatique de la Parole), partiellement financée dans le cadre du projet ANR ETAPE, a pour buts principaux de promouvoir une dynamique de l'évaluation en France, autour du traitement de la parole de la langue française, de mettre en place une structure pérenne d'évaluation et de diffuser le plus largement possible les informations et les ressources concernées par ces évaluations. La campagne ETAPE commence le 1^{er} juillet 2011 et se termine en mai 2012. Dans le cadre de cette campagne d'évaluation, ELDA et la DGA ont annoté des données télévision et radio qui sont mises à disposition des participants dans le respect des règles établies par le présent contrat. Par ailleurs, le CNRS met à disposition des participants la partie annotée en entités nommées du corpus ESTER (ELRA-S0241) sous la forme du « Corpus QUAERO d'émissions radiophoniques annoté en entités nommées ». Ce corpus ne peut être distribué sans l'obtention des transcriptions orthographiques du corpus ESTER, lequel corpus est fourni gracieusement par ELDA aux participants académiques d'ETAPE.

IL A ÉTÉ CONVENU ET ARRÊTÉ CE QUI SUIT :

"Les parties" désigne ELDA et L'UTILISATEUR.

1. Les **ressources linguistiques**, objet de ce contrat, pour lesquelles ELDA dispose d'un droit de distribution cédé par l'ayant-droit (ci-dessous le FOURNISSEUR) pour leur utilisation dans le cadre de la campagne d'évaluation ETAPE, sont définies en annexe A.

2. Le lieu d'utilisation des **ressources linguistiques** est indiqué en annexe B.
3. ELDA cède, à titre non exclusif, à l'UTILISATEUR, qui accepte, le droit d'utiliser les **ressources linguistiques** dans le cadre exclusif de l'évaluation de systèmes pour la campagne ETAPE.
4. ELDA concède à l'utilisateur, à titre gracieux et non-exclusif, le droit d'utiliser **les ressources linguistiques**, objet de ce contrat, dans le cadre exclusif de la campagne ETAPE et ensuite, **(et selon les termes de l'article 5)** dans le cadre exclusif de ses activités de recherche en ingénierie linguistique.
5. Par ce présent contrat, l'UTILISATEUR s'engage à participer aux activités d'évaluation de systèmes dans le cadre de la campagne ETAPE et à participer à au moins une des tâches définies par le plan d'évaluation et listée en annexe C. En cas de non soumission de résultats, l'UTILISATEUR perd son droit à conserver les **ressources linguistiques** et s'engage à les retourner à ELDA et/ou à effacer toute copie reçue.
6. L'UTILISATEUR n'est pas autorisé à reproduire les **ressources linguistiques**, ni à distribuer des produits dérivés ou services incluant tout ou partie des données dans un but commercial ou non (distribution gracieuse), sous toute forme ou par tout moyen, qui permettraient de reconstituer les **ressources linguistiques** ou une partie des **ressources linguistiques**. Les résumés, analyses et interprétations des propriétés linguistiques de l'information peuvent être dérivés et publiés à condition de ne pas rendre possible la reconstitution de l'information à partir de ces résumés. Des échantillons de l'information peuvent être montrés aux tiers ou publiés dans un cadre scientifique ou technique uniquement et dans le seul but de décrire les enjeux liés à la recherche et au développement. Un tel usage devra se faire dans le respect des droits des tiers, notamment des auteurs et des éditeurs.
7. ELDA accorde à l'UTILISATEUR le droit de reproduction, temporaire ou permanent, de traduction, d'adaptation et de modification, par tout moyen, des **ressources linguistiques**, si tout ou partie de ces actes sont nécessaires pour accéder et utiliser le contenu des **ressources linguistiques** de la manière convenue. L'UTILISATEUR est autorisé à dupliquer les **ressources linguistiques** uniquement dans un but d'archivage.
8. La concession de droit d'utilisation dont bénéficie l'UTILISATEUR n'entraîne le transfert d'aucun droit de propriété sur tout ou partie des **ressources linguistiques**.
9. L'UTILISATEUR reconnaît accepter les **ressources linguistiques** "telles quelles" c'est-à-dire avec tous les défauts qui peuvent subsister, et sans aucune garantie de l'adéquation de telles ressources pour un usage particulier.
10. L'UTILISATEUR s'engage à faire référence à ELDA, la DGA, ETAPE et au FOURNISSEUR dans ses publications qui mentionnent les **ressources linguistiques**.
11. L'UTILISATEUR s'engage à ne pas utiliser le nom d'ELDA, de la DGA, d'ETAPE ou du FOURNISSEUR pour la promotion commerciale de ces produits ou services. Il s'engage à éviter toute publication qui laisserait sous-entendre une approbation ou l'aval d'ELDA, de la DGA ou du FOURNISSEUR pour lesdits produits ou services.
12. Les parties ne peuvent être tenues pour responsables vis-à-vis l'une de l'autre des dommages, directs ou indirects, tels que la perte de bénéfices, les pertes de commandes, préjudices financier ou commercial qui résulteraient de l'exécution de ce contrat.
13. Les parties ne peuvent être tenues pour responsables de l'inexécution de leurs obligations prévues dans le présent contrat si celle-ci est due à un cas de force majeure ou bien à des circonstances qui échappent raisonnablement à leur contrôle.
14. L'UTILISATEUR s'engage à maintenir et publier une liste des personnes qui ont et qui ont eu accès aux **ressources linguistiques**. ELDA peut à tout moment et dans un délai raisonnable se réserver le droit de contrôler ces listes, en faisant appel à un auditeur indépendant de son choix, lequel s'engage à respecter la confidentialité des détails des affaires commerciales de l'UTILISATEUR.
15. Chacune des deux parties s'engage à garder confidentiels, les données, renseignements et documents divers (et en particulier, ceux de nature commerciale) communiqués par l'une ou l'autre des parties ou

dont elle aurait pris connaissance à l'occasion de l'exécution du présent contrat. Chacune des parties s'engage à faire respecter ces dispositions par ses agents et employés.

16. Les organisateurs du projet ETAPE ont établi un ensemble de directives limitant la diffusion et la publication des résultats de la campagne. Ces directives ont pour but de prévenir la publication d'informations inexacts ou incomplètes qui pourraient nuire à la réputation de la campagne ou des participants. L'UTILISATEUR s'engage à respecter les directives de publication définies en Annexe D.

17. Le présent contrat est régi par le droit français. Les parties déclarent leur intention de chercher une solution amiable à toute difficulté qui pourrait surgir à propos du contrat. En cas d'impossibilité et à défaut d'accord entre les parties, le Tribunal de Commerce de Paris sera le seul compétent pour connaître du litige. ELDA se réserve le droit d'annuler le présent contrat en cas de non-respect par l'UTILISATEUR d'un ou plusieurs articles du présent contrat, y compris l'article 16. Si l'une des clauses du contrat est nulle au regard d'une règle de droit ou d'une loi en vigueur, elle sera réputée non écrite mais n'entraînera pas la nullité de l'ensemble du contrat.

Le présent contrat comprend 17 articles ci-dessus, ainsi que les annexes A, B, C et D ci-dessous.

Signatures des parties

Fait à, en 2 exemplaires

Représentant de l'UTILISATEUR

Nom :

Titre :

Date :

Au nom d'ELDA

Nom : Khalid CHOUKRI

Titre : PDG d'ELDA

Date :

ANNEXES

ANNEXE A Description des ressources linguistiques, objet du contrat :

Corpus d'émissions télévisées

BFM: 7 heures, septembre 2010 – juin 2011, Copyright BFM. Droits de distribution : ELDA (Agence pour l'Evaluation et la Distribution des Ressources Linguistiques)

LCP – Assemblée nationale: 19 heures, septembre 2010 – juin 2011, Copyright « LCP – Assemblée nationale ». Droits de distribution : ELDA (Agence pour l'Evaluation et la Distribution des Ressources Linguistiques)

TV8 Mont-Blanc: 2 heures, février 2011 – juin 2011, Copyright TV8 Mont-Blanc. Droits de distribution : ELDA (Agence pour l'Evaluation et la Distribution des Ressources Linguistiques)

Corpus d'émissions radiophoniques

France Inter: 14 heures enregistrées et transcrites par la Direction Générale de l'Armement (DGA). Droits de distribution : ELDA (Agence pour l'Evaluation et la Distribution des Ressources Linguistiques)

Corpus QUAERO d'émissions radiophoniques annoté en entités nommées :

188 fichiers de transcriptions manuelles issus du corpus ESTER, contenant 1 291 225 mots incluant 113 885 types et 146 405 composants annotés.

Copyright : ELDA et CNRS

Droits de distribution : ELDA (Agence pour l'Evaluation et la Distribution des Ressources Linguistiques)

NB1 : Le corpus QUAERO d'émissions radiophoniques annoté en entités nommées ne peut être distribué sans le corpus ESTER (au minimum la partie « transcription orthographiques »), référencé ELRA-S0241 dans le catalogue ELRA (<http://catalogue.elra.info>).

NB2 : Le corpus ESTER (référencé ELRA-S0241 dans le catalogue ELRA (<http://catalogue.elra.info>)) est fourni gracieusement par ELDA aux participants académiques de la Campagne ETAPE.

ANNEXE B Lieu d'utilisation : (veuillez insérer ci-dessous l'adresse du lieu où les données seront exploitées)

ANNEXE C Description des tâches à effectuer par l'UTILISATEUR

- Transcription (T):
 - TRS - Transcription orthographique
- Segmentation/Suivi de Locuteurs (S):
 - SES-2 - Suivi des locuteurs multiples
 - SRL - Segmentation et regroupement de locuteurs
 - SRL-X - Segmentation croisée et regroupement de locuteurs
- Entités Nommées (E):
 - EN-ref - Détection des entités nommées dans la transcription de référence
 - EN-asr - Détection des entités nommées dans des transcriptions automatiques

ANNEXE D Conditions d'utilisation des publications et dissémination des résultats dans le projet ETAPE sont les suivantes:

- α) **PUBLICATION DES RESULTATS PAR LES ORGANISATEURS** : L'utilisateur accepte que les performances mesurées dans le cadre des tests officiels de la campagne d'évaluation ETAPE et validées lors de l'atelier de clôture de la campagne soient publiées de manière non anonyme par les organisateurs. Ces publications se conformeront aux règles générales de la présente annexe. Les projets publications seront soumis aux participants et les remarques prises en compte.
- β) **PUBLICATIONS DES PARTICIPANTS** : Dans la mesure où il respecte les règles générales de la présente annexe, l'utilisateur reste libre de publier ses propres résultats. Il peut aussi librement faire référence aux publications des organisateurs. Il peut reprendre partiellement ces publications dans la mesure où il respecte les règles d'impartialité et d'objectivité. En dehors de ces cas, il ne peut citer les résultats des autres participants sans leur accord.
- γ) **REGLES GENERALES POUR LES PUBLICATIONS SCIENTIFIQUES OU TECHNIQUES** : Les publications scientifiques ou techniques, notamment les lettres d'informations universitaires ou de laboratoires de recherche, doivent respecter les normes d'impartialité et d'objectivité établies par la communauté et énoncer d'une manière claire et précise les limites des conditions de test, ainsi que les autres facteurs pouvant influencer les résultats. La nature expérimentale des tâches, des données et des procédures d'évaluation doit également être précisées.
- δ) **PUBLICITE ET COMMUNIQUE DE PRESSE** : Aucune annonce publicitaire utilisant les résultats d'évaluation d'ETAPE ne peut être publiée dans les journaux, périodiques, revues, ou autres publications. Les communiqués de presse sur les résultats d'ETAPE auprès d'organismes ayant une couverture nationale et internationale sont également interdits.
- ε) **PUBLICATIONS MARKETING, LETTRES D'INFORMATIONS LOCALES**: Bien qu'il soit admis que des discussions étendues sur le thème de l'évaluation ne sont pas appropriées dans ce type de publications, toute affirmation établie sur la base des résultats d'évaluation doit être fondée. Les mesures d'évaluation utilisées pour justifier ces affirmations doivent être clairement présentées et les actes de conférence référencés. Là où le matériel promotionnel fait l'objet d'une révision avant publication par les médias, l'auteur doit contrôler que la révision ne cause aucune violation aux présentes directives.