

LNE-Visu : une plateforme d'exploration et de visualisation de données d'évaluation

Guillaume Bernard¹, Olivier Galibert¹, Rémi Régnier¹, Séverine Demeyer¹, Juliette Kahn¹
(1) Laboratoire National de métrologie et d'Essais, 29 rue Roger Hennequin, 78190
Trappes, France
prenom.nom@lne.fr

RESUME

LNE-Visu est une interface de visualisation et d'exploration de données multimédia qui regroupe les données des campagnes d'évaluation organisées par le LNE. Elle propose 3 fonctionnalités principales : explorer et sélectionner des corpus, visualiser et écouter des données et effectuer des tests de significativités de différences.

ABSTRACT

LNE-Visu : a tool to explore and visualize multimedia data

LNE-Visu is a tool to explore and visualize multimedia data created for the LNE evaluation campaigns. 3 functionalities are available: explore and select data, visualize and listen data, apply significance tests

MOTS-CLES : Visualisation de données multimédia, Campagnes d'évaluations, Significativité, MAURDOR, TRAD, REPERE, FABIOLÉ, ETAPE.

KEYWORDS: Multimedia data visualization, Evaluation campaigns, significance, MAURDOR, TRAD, REPERE, FABIOLÉ, ETAPE.

1 Contexte

La communauté du traitement automatique de la langue écrite et parlée s'appuie sur des campagnes d'évaluations pour valider les approches développées. Lors des campagnes d'évaluation, les hypothèses de différents systèmes sont comparées simultanément sur une tâche précise (transcription de la parole, détection d'entités nommées, traduction automatique...) à partir des mêmes données inédites et avec les mêmes métriques. Le LNE, en tant que tiers de confiance pour des organismes publics, organise ou co-organise des campagnes d'évaluation pour des systèmes de traitement automatique de données multimédia (parole, texte, image, vidéo...) comme ETAPE (Galibert et al, 2014), REPERE (Galibert et Kahn, 2013), (Galibert et al, 2014), MAURDOR (Oparin et al, 2014), (Galibert et al, 2014), TRAD (Galibert et al, 2013), QUAERO (QUAERO, 2010) ... Ces campagnes ont donné lieu à la création des outils d'évaluation LNE-tools et au développement d'une plateforme de visualisation de données d'évaluation LNE-visu.

En effet, évaluer des systèmes nécessite de bien connaître les données mises à la disposition des développeurs (données d'apprentissage, de développement et de test) notamment pour mettre en évidence les facteurs de variation de la performance. Cette connaissance passe par la visualisation

ou l'écoute des données de manière détaillée. Obtenir également des visualisations synthétiques est une nécessité pour pouvoir comparer. Par ailleurs, la création des données étant très coûteuse, il est intéressant de pouvoir mutualiser les annotations de différents corpus afin de pouvoir les comparer indépendamment de leur origine de campagne. De nombreuses initiatives (ISIDORE (Huma-Num, 2010), COCOON (Jakobson et al, 2012), ELDA/ELRA (ELDA, 2016)...) ont vu le jour pour mutualiser les données en se basant principalement sur une nomenclature commune de métadonnées globales (TEI, plateforme CLARIN (Caillau et al, 2010)...) permettant de faciliter l'indexation. Elles ne permettent pas cependant d'explorer précisément les corpus avec un outil unifié. La plateforme LNE-visu cherche à répondre aux défis suivants :

- Exploration de corpus existants,
- Sélection de données de corpus,
- Visualisation synthétique des données choisies,
- Visualisation des données selon différentes variables et facteurs,
- Application de traitements statistiques.

2 Fonctionnalités de la plateforme de visualisation de données d'évaluation

2.1 Explorer et sélectionner des corpus

Les corpus pouvant actuellement être explorés via LNE-visu sont les corpus REPERE, QUAERO Entités nommées, ESTER 1 et 2. Les autres corpus (MAURDOR, TRAD...) à la disposition du LNE seront intégrés dans l'outil dans les prochains mois. Il est possible de sélectionner via une data-frame tout ou partie de ces corpus afin de visualiser et écouter les données.

2.2 Visualiser et écouter des données

Les données peuvent être visualisées et écoutées. Il est également possible d'obtenir des synthèses des informations présentes (Figure 1).



Figure 1: Exemples de visualisation

2.3 Effectuer des tests de significativité de différences

Des tests de significativités de différences (ANOVA, t.test) ou de corrélation (Spearman, Kendall...) sont accessibles via la plateforme.

Références

- GALIBERT O., LEIXA J., ADDA G., CHOUKRI K., ET GRAVIER G., (2014), « The ETAPE speech processing evaluation », Actes de *LREC 2014*.
- GALIBERT O ET KAHN J, (2013), « The First Official REPERE Evaluation », Actes de *First Workshop Speech Lang. Audio Multimed. SLAM 2013*.
- BERNARD G, GALIBERT O, KAHN J, (2014), « The second official REPERE evaluation », Actes de *SLAM 2014-Second Workshop on Speech, Language and Audio for Multimedia*.
- OPARIN I, KAHN J, ET GALIBERT O, (2014), « First Maurdor 2013 Evaluation Campaign in Scanned Document Image Processing », Actes de *IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP'14*.
- GALIBERT O, KAHN J, ET OPARIN I, (2014), « Results of the Second MAURDOR _evaluation campaign », présenté à Atelier de restitution des résultats de la seconde campagne MAURDOR, Paris
- GALIBERT O, KAHN J, OPARIN I, ET TABARY F, (2012), « TRAD 2012 _Evaluation Campaign : Results », présenté à Atelier de restitution des résultats de la campagne TRAD 2012, Paris, nov-2012.
- QUAERO, (2014), « Quaero - Programme de recherche industrielle sur les technologies d'analyse, de classification et d'utilisation de documents », *Quaero*. [En ligne]. Disponible sur: <http://www.quaero.org/>. [Consulté le: 03-févr-2014].
- HUMA-NUM, (2010), « ISIDORE, une plateforme de recherche pour les données des SHS », in *2ème université du TGE Adonis*.
- JACOBSON M, BADIN F, ET GUILLAUME S, (2015), « Cocoon une plateforme pour la conservation et la diffusion de ressources orales en sciences humaines et sociales », Actes de *8es Journées Internationales de Linguistique de Corpus*.
- ELDA, « Evaluations and Language resources Distribution Agency ». [En ligne]. Disponible sur: <http://www.elra.info/en/>.
- CAILLIAU F, HINRICHS E, ET KRAUWER S, (2014), « The CLARIN Research Infrastructure: Resources and Tools for E-Humanities Scholars », Actes de *LREC 2014*.