

*LUE IMPACT OLKi - Open Language and Knowledge for Citizens

*Compte-rendu de la réunion #3 du groupe de travail sur la *plateforme OLKi*
13 juin 2019 de 15h à 16h30 à la MSH, salle 326

Documents de travail

Dossier B'UL regroupe les documents: <https://bul.univ-lorraine.fr/index.php/s/K5gWr1HljdonFvZ>

Relevé des tâches à faire

	Quoi ?	Qui ?	Pour quand ?
1	Contribuer à l'article <ul style="list-style-type: none"> - affiner si besoin les objectifs, cibles, forme, etc - proposer une revue de destination - proposer un plan - recenser des références bibliographiques 	Tous	A définir
2	Proposer des améliorations aux schémas ou des schémas alternatifs selon les points de vue, les cas d'usage	Tous	A définir

Compte-rendu

1. Nouveauté de la plateforme

Nouveauté :

Lien vers l'instance de démonstration : <https://olki-social.loria.fr/fr>

- Front end web (instance de démonstration): <https://olki-social.loria.fr/fr>

Affichage avec métadonnées minimales. Vérifiez si Dublin Core.

- Back end (python) :
 - o gestion des comptes utilisateurs (login, sign up)
 - o gestion des corpus (= un ou plusieurs fichiers texte ou autre),
 - o à venir, les interactions avec ActivityPub et le scrapping
- API ouverte utilisée est celle autre que REST

Idée de faire un système de plugins pour moissonner un entrepôt par ex. ArXiv.

Idée à creuser : Formaliser la constitution d'un comité technique ? Objectifs à énoncer.

2. Réflexions, pistes de développement

Problématique de la granularité de l'accès à l'intérieur du corpus *ie* à des objets et sous-objets

Dans un premier temps, volonté de rester général dans les documents qui peuvent être utilisés, mais toutefois difficultés pratiques (voire théoriques?) pour permettre la prise en charge de formats de façon généraliste. D'où la nécessité d'avoir un système de plugins par format de données.

Idée de développer les outils (pointeur, ...) dans un deuxième temps pour travailler sur une partie de texte par exemple.

Ex des lettres de Poincaré

Corpus → l'ensemble des lettres ; nécessité d'avoir un fichier de métadonnées par document à l'intérieur du corpus.

Notion de ressource : 1 fichier/1 item = 1 ressource ; une collection de ressources = 1 corpus

➔ Métadonnées à quel niveau ?

- ORTOLANG: métadonnées au niveau de chaque ressource
- OLKi : métadonnées également au niveau de chaque ressource, il faudrait envisager d'avoir des métadonnées aussi au niveau du corpus/collection

Solution généraliste peu satisfaisante. Option des plugins semble plus pertinente.

Plateforme OLKi :

- communication autour des ressources et non pas hébergement sur le long terme des données.
- accès aux métadonnées dans d'autres réseaux
- pas de corpus de tweets car problèmes juridiques

Pour l'instant : outil de recherche simple

Si projet ANR avec ORTOLANG est accepté : lemmatisation,

WM : projet TEI de faire une plateforme de crowdsourcing autour de ressources en polonais

Source physique (auteur) ➔ source numérique (contributeur fait entrer le document numérique dans le réseau)

Idée de développer un plugin interagissant avec les sites OMEKA et NAKALA.

Pistes à explorer pour les fonctionnalités

- Google Scholar et Google Dataset : lien vers les articles scientifiques citant les corpus, lien entre articles , ...sont des plus-values scientifiques.

<https://toolbox.google.com/datasetsearch>

<https://www.google.com/publicdata/directory#!>

- Ulysse (service UL d'accès aux ressources des bibliothèques de l'UL)

Outil de veille scientifique : un stagiaire de CC travaille sur le développement d'un outil de veille qui doit pouvoir trouver des articles similaires selon une « distance sémantique » intéressante à trouver.

Attention à bien faire apparaître le lieu de stockage du corpus/l'origine de la ressource : réflexion en cours sur le format de différenciation (couleur, colonne locale vs moissonnée)

Proposer des rôles pour les contributeurs :

Pistes à explorer pour le design de l'interface :

- stagiaire
- professionnelle freelance en Pologne (contact de WM)
- projet d'étudiants en design : explorer côté ARTEM ou formation informatique avec spécialité design, école de jeux vidéos à Strasbourg (<http://ludus-academie.fr/>)

3. Communication

Attention à la terminologie de présentation/éléments de langage : « réseau de communication interdisciplinaire » vs « réseau social » (connoté négativement).

Problème actuel de cadrage et de communication.

Idées de communication : démo, use cases en vidéo

- *use case* sur les transcriptions de manuscrit quand il y a un doute sur un mot (car trop mal écrit par exemple)
- *use case* orienté sur la médiation scientifique en cherchant le meilleur mot. Des médiateurs des sciences (« vulgarisateurs ») pourrait investiguer un mot en particulier.

Un plan de communication pour la plateforme est proposé par PAR, CC, et AC. La direction du Numérique pourra être contactée.

4. Positionnement de la plateforme OLKi

Nécessité de clarifier le périmètre de la plateforme.

Un schéma des acteurs, des ressources, des outils et des flux d'information est proposé mais n'a pas été discuté.

Un schéma positionnant la plateforme (ses services) dans le cycle de la donnée et les services UL proposés aux personnels et étudiants est proposé.

- ➔ Attention vision « gestionnaire de données », nécessité de faire le schéma sous les angles veille bibliographique, corpus, réseau social/communication, etc...

DARIAH : plateforme de communication pour mettre en relation les gens, les documents, les compétences à l'échelle européenne.

Nécessité de positionner OLKi par rapport à DARIAH, Dataverse, ...Google ?

5. Article

Sujet non abordé. Projet en attente du choix d'une revue.

Lien vers le calepin : <http://etherpadlite.univ-lorraine.fr/p/l58vfy90op>

Lien vers la forge logicielle: <https://framagit.org/synalp/olki/documents/seminar-article>

6. Nom de la plateforme

Propositions de JB

- MuseFed (« Média Universel pour le Suivi et l'Édition Fédérative », mais un peu grandiloquent peut-être)
- SciFed (plus généraliste)
- SocialKi (prononcer à l'anglaise, permet de garder le beau logo de feuille)
- SciShare (un peu facile)
- FediScience
- FediShare

7. Fonctionnement du groupe

Les membres présents relisent le compte-rendu et communiquent les modifications au rédacteur. S'il n'y a pas de modification à apporter, le compte-rendu d'une réunion est approuvé 3 jours avant la réunion suivante. La version anonyme sera mise en ligne 2 jours avant la réunion suivante.

Prochaine rencontre

Date à définir

Ordre du jour prévisionnel :

- Approbation/corrections du précédent compte-rendu (5')
- Présentation des nouveautés (10')
- Travail sur les schémas (30')

- Choix d'une revue pour l'article (30')

Rédactrice : Aurore Coince

Pour rejoindre la visioconférence

- connectez-vous sur <https://visio.inria.fr> (<login>@inria.fr) puis "Se réunir" : gt.platforme@visio.inria.fr
- pour les externes <https://visio.inria.fr/invited.sf?secret=7OJkn78ZSd2mT0hjm0gBOg&id=130784964>
- depuis un terminal de visioconférence : appelez visio.inria.fr (IP: 128.93.162.102) puis entrez le numéro de la conférence : 130784964#
- par téléphone : +33 4 92 38 77 88 (77788), puis entrez 130784964#