

\*LUE IMPACT OLKi - Open Language and Knowledge for Citizens

\*Compte-rendu de la réunion du GT plateforme OLKi du 2 mai 2019 de 16h à 17h30 au Loria, salle B009

### Documents de travail

- <https://framagit.org/synalp/olki/documents/software-requirements-specification/releases>
- <http://etherpadlite.univ-lorraine.fr/p/4tmt dih7eo>
- <https://olki.loria.fr/blog/usecase/>

**Objectifs de la réunion** : présenter le projet et les acteurs ainsi que se mettre d'accord sur les objectifs et le mode de fonctionnement du GT.

### Agenda (1h30)

- Présentation du projet de plateforme (15')
- Discussion autour des objectifs de la plateforme (10')
- Point sur les objectifs et le mode de fonctionnement du GT du GT (10')
- Tour de table (20') : présentation, affiliation, contributions/apports et gains

### Compte-rendu

#### 1. Projet de plateforme

Objectif de la plateforme : créer un écosystème ouvert et fédéré favorisant le travail et les échanges scientifiques autour de ressources langagières notamment

Principes de la plateforme:

- plateforme ouverte à tous soumise à la création un compte utilisateur pour mettre en ligne un corpus ou échanger sur la plateforme
- plateforme interconnectée avec le Fediverse, fédération promouvant une communauté locale structurée, autour de corpus par exemple.

Fonctionnalités :

- Mise en ligne de ressources (hébergement de corpus structurés ou non), référencements, contributions et annotations, connexion à autres plateformes ...
- Mise à disposition d'outils de visualisation et de recherche à l'intérieur de corpus.
- Communication (échanges entre scientifique mais aussi avec des citoyens via les réseaux sociaux existants)
- Veille scientifique (ex. moissonnage d'ArXiv par un robot)

Non-fonctionnalités :

- pas d'archivage mais complémentarité avec les autres plateformes (ORTOLANG qui permet la conservation sur le long terme, pérennisation des ressources, Zenodo, ...)
- pas de création/constitution de corpus directement sur la plateforme mais réutilisation de corpus existants et évolution de corpus
- pas de formalisation des métadonnées, pas d'indexation de l'intérieur d'un corpus mais indexation très simple (titre et description d'un corpus en texte brut)

Quel format pour les commentaires ? transfert des commentaires/des images ?

ActivityPub est un standard, format JSON Id des commentaires

Standard Web annotation

Comment le citoyen lambda va s'y retrouver ?

But est de communiquer : poser des questions notamment et de créer une communauté autour du responsable du corpus.

Démonstrateur interne prévu pour sept 2019.

### Pistes de travail

- Réfléchir à la finalité de la plateforme : objectif qualitatif ou quantitatif ?
  - o Outils différents à développer selon la finalité envisagée
- Proposer un positionnement d'OLKi au sein du mouvement de la Science Ouverte, et explicitation/clarification par rapport à l'existant. Thomas fera une proposition de positionnement d'OLKi dans le cycle de vie de la donnée.
- Interfaçage avec le projet VisaTM à envisager. VistaTM propose aussi des échanges entre chercheurs : <https://visatm.inist.fr/>
- Nécessité d'imaginer la valeur ajoutée, nécessité de formuler un cas d'usage pour faire le travail de veille.
- Réflexion à mener sur le caractère flexible ou contraignant de l'interface
  - o Interface contraignante permet de rassembler des corpus de provenances différentes.
- Réfléchir au système de plugins
  - o Envisager de créer un plugin permettant le dépôt vers des entrepôts institutionnels (Dataverse, Zenodo, ORTOLANG...) recommandés par le CoSO.  
RQ : si plus de cibles, plus de travail à faire pour développer des plugins
  - o Possibilité d'intégrer des plugins spécifiques de traitement à la plateforme, etc...format du plugin sera dépendant des ressources et de ses objectifs.
  - o Permettre à chacun de développer ses plugins particuliers
- Idée de rédiger un article fondateur

### Informations diverses

- Consortium CORLI (Corpus, Langues, Interactions) (<https://corli.huma-num.fr/new/drupal/>) lance un appel sur la finalisation de corpus visant à être intégré à des entrepôts (ORTOLANG)
- Voyant tools : ensemble d'outils d'exploitation de corpus (non-structuré) sur du texte (<https://voyant-tools.org/?corpus=shakespeare>)
- Questionnaire sur la fouille de textes par l'INIST : <https://www.inist.fr/nos-actualites/questionnaire-fouille-de-textes/> (les résultats devraient être publiés à l'automne)

## **2. Objectifs et mode de fonctionnement du Groupe de Travail**

### **a. Objectifs**

Le GT "plateforme OLKi" a pour objectif d'orienter le développement de la plateforme fédérée OLKi afin qu'elle réponde aux besoins des cibles (scientifiques en premier lieu), que les cibles se l'approprient et que le modèle se diffuse, en cohérence avec l'existant mais aussi avec ce qui se développe en parallèle.

Concrètement, les activités proposées au groupe sont:

- signaler les points forts et les points faibles du projet de plateforme (objectifs, méthodologie, fonctionnalités, planning, ...)

- partager les pratiques scientifiques de communication autour des ressources langagières notamment
- faire la veille sur les plateformes existantes et les initiatives qui pourraient être similaires ([EOSC](#), etc...)
- positionner le projet par rapport au paysage international concernant les données de recherche et leur ouverture
- guider l'implémentation des fonctionnalités suivant le respect des standards ([DataCite](#), données FAIR, etc.)
- faire de la veille sur les fonctionnalités proposées par les plateformes éditoriales (en termes de conversation scientifique)
- proposer des idées de développement et de fonctionnalités
- réfléchir à la meilleure utilisation des ressources disponibles (humaines et financières) ainsi qu'aux pistes pour augmenter ces ressources (Appel ANR Flash Open Science, programme Etalab des Entrepreneur d'Intérêt Général, etc...)
- participer à la production de documents ou contenu (rédaction et/ou relecture d'un article fondateur, des spécifications, ...)
- faire des retours d'expérience
- construire un réseau de contributeurs et d'utilisateurs de la plateforme
- réfléchir à et faire la promotion de la plateforme

## b. Mode de fonctionnement

Prochaines réunions en présentiel avec visioconférence.

Pour rejoindre :

- connectez-vous sur <https://visio.inria.fr> (<login>@inria.fr) puis "Se réunir" : gt.plateforme@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 puis entrez le numéro de la conférence : 130784964#
- par téléphone : +33 4 92 38 77 88 (77788), puis entrez 130784964#

Un dossier B'UL regroupe les documents utiles au groupe :

<https://bul.univ-lorraine.fr/index.php/s/K5gWr1HljdonFvZ>

Privilégier les schémas pour intégrer les personnes intéressées qui vont rejoindre le groupe au fur et à mesure. Ex : Mind map (attention généralement pas cyclique) avec lien vers la doc ou d'autres schémas plus spécifiques.

## 3. Tour de table

J. B., étudiant en physique, étudiant-entrepreneur, intérêts pour le machine learning, la science reproductible, le logiciel libre, programmeur. Observateur au début.

A. B., équipe Multispeech du Loria, phonéticienne, travail sur les indices acoustiques, interférences entre langues premières et langues secondes, intérêts pour échanger des connaissances entre linguistes et philosophes, intérêts pour les échanges autour des connaissances (par rapport aux corpus)

M. C., enseignant-chercheur en informatique à l'ATILE, spécialisé dans la construction de corpus linguistiques, traitement automatique des langues, membre du comité de pilotage du projet OLKi, intérêt pour faire le lien avec ORTOLANG

C. C., chercheur CNRS, porteur scientifique du projet OLKi, souhaite l'opinion du groupe de travail, exploration des usages, des besoins de la communauté scientifique pour faire évoluer à la marge ou plus la plateforme.

Aurore Coince, Project Manager du projet OLKi, gère le projet dans son ensemble, coordonne les acteurs, suit les activités et le budget, organise la communication et les événements.

J. F., chercheur associé au CREM, porteur de projet distribué fédéré Needle. Curiosité. Valeurs similaires entre les projets OLKi et Needle, synergie à trouver.

S.G., enseignant-chercheur associé au CREM en info-comm, spécialités en veille stratégique, culture en information et linguistique. Intérêt pour connaître la plateforme.

T. J., Dir de la doc scientifique, membre du comité de pilotage du groupe de travail science ouverte (objectif de renforcer le volet opérationnel de la science ouverte). Intérêt pour apporter de la complémentarité avec les nouvelles réglementations sur l'ouverture des données, par ex. interaction avec les archives pérennes.

Pierre-Antoine Rault, ingénieur responsable du développement

P. W., ingénieur aux Archives Henri Poincaré, travail sur des archives, analyse de sources & humanités numériques, la valorisation par mise en ligne de corpus plus ou moins structurés, et la création d'une ontologie pour les métadonnées alignées sur des ontologies existantes.

Rédactrice: Aurore Coince ([aurore.coince@univ-lorraine.fr](mailto:aurore.coince@univ-lorraine.fr))