

MARDI 22 OCTOBRE 2024, 15H-17H

Présences

Anaïs Chambat, Élysabeth Hue-Gay, Marie-Agnès Lucas-Avenel, Emmanuelle Morlock, Marianne Reboul, Nathalie Rousseau

Ordre du jour

- Bilan de l'atelier pour 2024
- Objectifs de l'atelier pour 2025
- Questions diverses

Discussions

- Présentation des travaux effectués par l'atelier « documents lexicographiques » :

L'un des objectifs du GT3 est de réaliser un recensement des outils utilisés par des projets existants, afin que les collègues désireux de se lancer puissent avoir une idée de la diversité des outils.

Ce recensement sera publié sous la forme d'un Scenari, dont l'atelier « textes lexicographiques » élabore en ce moment un prototype :

<https://axe-1-gt3-outils-et-pratiques-editoriales.gitpages.huma-num.fr/lex/recensement>

compte rendu de la réunion de l'atelier « documents lexicographiques » (lien provisoire) : <https://test-ariane-1907-2-afall4-05c45eb02e82f03fcbce2546ef50467de43a9.gitpages.huma-num.fr/layers/lex.html>

- Présentation des travaux effectués par Lucie Cordier durant son stage de 140 h qui s'est déroulé au cours de l'année 2024 : questionnaire visant à recenser les différents outils utilisés et/ou créés dans les projets d'édition numérique de corpus en langues anciennes, réalisés à partir de sources manuscrites et/ou épigraphiques. Il reste à retravailler la partie portant sur les outils pour l'adapter au mieux au Scenari existant (et inversement), N.B. : Dans Scenari, il est possible de placer une partie des informations dans des champs libres, et une autre partie dans des critères qui peuvent être utilisés pour trier les fiches. Voir la version actuelle du recensement de scripts élaboré par l'atelier scripts, qui a servi de modèle : [https://recensement-axe-1-gt3-outils-et-pratiques-editor-01966b4a0dba95.gitpages.huma-num.fr/?state=\[\]&tab=keys](https://recensement-axe-1-gt3-outils-et-pratiques-editor-01966b4a0dba95.gitpages.huma-num.fr/?state=[]&tab=keys)
- Pascale Paré-Rey porte un projet de dispositif national pour contribuer à la structuration de la communauté scientifique étudiant les textes anciens qui démarre actuellement. Il conviendrait de la contacter pour voir quelles synergies sont possibles avec notre atelier.
- Il conviendrait de réfléchir, à moyen terme, à la diffusion des résultats du questionnaire de notre atelier ; de manière générale, penser à DARIAH

Décisions

- L'atelier se fixe comme objectif, pour 2025, de publier le questionnaire après avoir retravaillé la partie portant sur les outils, et d'en publier les résultats dans Scenari
- NR va prendre contact avec Pascale Paré-Rey
- NR va relayer auprès des membres du consortium l'information concernant la possibilité de publier des workflows sur <https://marketplace.sshopencloud.eu/>
- Prochaine réunion le 10 décembre 2024, 14h-16h

MARDI 10 DÉCEMBRE 2024, 14H-16H

Présences

Anaïs Chambat, Marie-Agnès Lucas-Avenel, Emmanuelle Morlock, Pascale Paré-Rey, Nathalie Rousseau

Ordre du jour

- Bilan des actions effectuées depuis la dernière réunion
- Présentation du **dispositif « Sciences des textes anciens »** lancé à l'automne 2024 par le MESRI et coordonné par Pascale Paré-Rey ; discussion pour définir quelles sont les synergies possibles entre nos travaux
- **Travail sur le questionnaire**, en particulier sur la partie portant sur les outils: **il est nécessaire de remplacer les questions des sections intitulées « Logiciels et services web », « Fichiers » et « Modèles informatiques » par des questions dont les réponses puissent s'insérer dans un Scenari**, de préférence mutualisé avec le Scenari ébauché par l'atelier « documents lexicographiques ». Question: que peut-on reprendre et que faut-il ajouter aux critères proposés par ce Scenari existant?
 - état actuel du questionnaire : <https://framaforms.org/questionnaire-2025-ariane-gt3-outils-et-pratiques-editoriales-editions-numeriques-de-corpus-antiques>
 - état actuel du Scenari, d'après les critères retenus par l'atelier « documents lexicographiques » : [https://axe-1-gt3-outils-et-pratiques-editoriales.gitpages.huma-num.fr/lex/recensement/?state=\[\]&tab=keys](https://axe-1-gt3-outils-et-pratiques-editoriales.gitpages.huma-num.fr/lex/recensement/?state=[]&tab=keys)
 - N.B. première partie du questionnaire : réfléchir à une éventuelle taxonomie pour le Scenari (voir la proposition actuelle pour les ressources lexicographiques dans Scenari, onglet « parcourir la taxonomie »)
- Questions diverses

Bilan des actions

- Assemblée générale du consortium, le 26 novembre dernier à Aubervilliers
- NR et PPR ont pris contact ; PPR rejoint l'atelier ; NR sera membre du dispositif « Sciences des textes anciens »
- NR a relayé auprès des membres du consortium l'information concernant la possibilité de publier des workflows sur <https://marketplace.sshopencloud.eu/>

Dispositif « Sciences des textes anciens »

- 1^{er} livrable : enquête sur la construction des savoirs critiques, sur les pratiques et les besoins des chercheurs, notamment via les usages du numérique
- 2^e livrable : cartographie assez vaste: projets, outils (avec référence au questionnaire d'ARIANE et au guide « Chaînes d'acquisition, de traitement et de publication du texte » : <https://hal.science/hal-04734959>), structures, ressources existantes (dont Biblissima)...
- 3^e livrable : formation, information, en particulier pour les universités qui n'ont pas de grandes structures de recherche (sous la forme de conférences, d'école thématique, avec invitations de chercheurs étrangers)
- 4^e livrable : site de ressources, avec résultats de l'enquête et de la cartographie
- 5^e livrable : une structure (consortium, institut ?) pour rassembler les chercheurs en LC en France

Calendrier : initié en juin, notification des fonds il y a peu, budget engagé en 2024, recrutement d'ingénieur début février, exécution du budget sur 2025 mais possibilité de continuer en 2026

Points communs et différences entre les actions du dispositif et le questionnaire ARIANE :

- L'enquête du dispositif « Sciences des textes anciens » actuellement en préparation est destinée au public français et plus largement sur les pratiques et les besoins des chercheurs, notamment numériques, mais pas uniquement ; quant à la cartographie, elle portera sur les ressources, les projets, également notamment numériques, mais pas uniquement (« à l'ère du numérique »).
- Le questionnaire ARIANE se concentre sur les projets numériques, et plus précisément sur les outils.

Remarque sur Biblissima

Biblissima n'effectue pas d'enquête ; en revanche,

- appel à projets annuel
- possibilités de participation aux clusters (ouverts)
- cluster 5b sur l'édition critique : discussions sur les meilleurs outils par typologie textuelle, sur le schéma TEI par typologie textuelle

Discussion sur les critères actuellement dégagés par l'atelier « Documents lexicographiques »

[en gras, modifications apportées, simultanément enregistrées par AC dans le prototype Scenari :
[https://axe-1-gt3-outils-et-pratiques-editoriales.gitpages.huma-num.fr/lex/recensement/?state=\[\]&tab=keys](https://axe-1-gt3-outils-et-pratiques-editoriales.gitpages.huma-num.fr/lex/recensement/?state=[]&tab=keys)]

1. Numérisation d'une ressource physique [OK ainsi]

Oui/Non

2. Objectifs numériques et scientifiques [modifié]

- Indexation Lien avec le fac-similé de la source
- Plein-texte (relu ou non)
- Texte balisé
- Texte annoté (commentaires de l'éditeur)
- Texte critique (établi par l'éditeur à partir de différentes versions)
- Édition imprimable

3. Opérations réalisées et outils numériques associés [modifié]

- Segmentation **du document** (ex. algorithmes de layout analysis, ROI ...)
- Transcription/Reconnaissance du texte (ex. eScriptorium, Rescribe, Transkribus, ABBYY FineReader...)
- Post-correction (ex. Pyrrha...)
- Lemmatisation (ex. Pie-Extended...)
- Balisage (ex. GROBID-Dictionaries, Oxygen, **XMLMind XML Editor**...)
- Métadonnées (ex. Mynkl...)
- Conversion (ex. Odette, TEIGarage...)
- Composition (ex. LaTeX, InDesign, MaX, TEIPublisher...).
- **Mise à disposition en ligne** Disponibilisation (ex. **MaX**, TXM, TEIPublisher...)
- Validation et améliorations collaboratives (ex. plateforme TACT, **PLUCO dans XXE**...)

4. Fonctionnalités proposées [à revoir]

- Recherche avancée (ex. présence de filtres et de tris: par lemmes, dates, mots-clés, etc.)
- Mise en relation diachronique ou synchronique (ex. d'éditions d'un même ouvrage, de ressources de nature différente)

- Multimédia (ex. audio, vidéo, image)
- Exportation **des données de l'édition**
- Édition imprimable

5. Modalités de consultation et d'accès [à discuter]

- Consultation seulement. Accès fermé: Les données sont consultables via une interface soumise à des restrictions (ex. abonnement, enregistrement). Pas de téléchargement possible.
- Consultation et téléchargement uniquement. Accès semi-fermé: Les données sont accessibles pour consultation et téléchargement sous certaines conditions (ex. licence restrictive, abonnement, enregistrement).
- Consultation, téléchargement et réutilisation. Accès ouvert: Les données sont librement consultables, téléchargeables et réutilisables, conformément aux conditions d'une licence ouverte.

6. Langues [à discuter]

- Français (fra)
- Allemand (deu)
- Anglais (eng)
- Espagnol (esp)
- Grec ancien (grc)
- Italien (ita)
- Latin (lat)
- Portugais (prt)
- Autre (norme ISO-639-3). Pour les langues rares, voir la norme ISO-15924 ou le catalogue Glottolog.

Questions diverses

Échanges autour de la citabilité :

- Les presses universitaires de Caen valident à la fois les éditions papier et les éditions numériques
- Exemple à méditer : RIDE <https://ride.i-d-e.de/>

Décisions

- lors de la prochaine séance, poursuite du travail sur le questionnaire (voir ordre du jour du 10/12)
- NR va élaborer la partie du questionnaire correspondant aux critères 1 à 3
- MALA va faire le lien avec le Cluster 5b de Biblissima

Consortium ARIANE

Axe 1 – GT3 « Outils et pratiques éditoriales »

Atelier Éditions de textes antiques (ANT)

Comptes rendus des réunions de l'année 2024

- le questionnaire « Sciences des textes anciens » sera articulé avec le questionnaire ARIANE
- AC va écrire à Ariane Pinche pour vérifier que les critères de la section 3 coïncident au maximum avec ceux qui ont été élaborés dans le document intitulé « Chaînes d'acquisition, de traitement et de publication du texte » (<https://hal.science/hal-04734959>)
- prochaine réunion le 14/01 de 16h15 à 18h