



« Consortium-HN ARIANE »

Analyses, Recherches, Intelligence Artificielle et Nouvelles Éditions de textes
(2023-2027)



Axe 1 – GT 3 « Éditions numériques de qualité »
Coordination : Anaïs Chambat et Nathalie Rousseau



LABORATOIRE DE RECHERCHE · EA 7518
LEXIQUES · TEXTES · DISCOURS
DICTIONNAIRES
CENTRE JEAN PRUVOST



CERGY PARIS

UNIVERSITÉ



LETTRES
SORBONNE
UNIVERSITÉ



Ordre du jour

- Tour de table rapide: présentations
- Présentation du Consortium
 - Enjeux
 - Exigences
- Présentation du GT et de son organisation en ateliers
 - Plusieurs types d'activités possibles...
 - Sur plusieurs types de sujets : ateliers thématiques / transverses
 - Calendrier prévisionnel des premières activités proposées
- Informations pratiques
- Tour de table: vos propositions d'ateliers / de livrables / d'activités pour 2024 / pour une année ultérieure ?

Tour de table rapide : présentations

Anaïs Chambat

Doctorante contractuelle en Sciences du langage

Laboratoire Lexiques, Textes, Discours, Dictionnaires – Centre Jean Pruvost (EA 7518)

CY Cergy Paris Université

<https://achambat.github.io/>

Nathalie Rousseau

Maître de conférences

Sorbonne Université – Faculté des Lettres, UFR de Grec

UMR 8167 Orient et Méditerranée

Institut universitaire de France

<https://lettres.sorbonne-universite.fr/personnes/nathalie-rousseau>

www.iufrance.fr/les-membres-de-liuf/membre/2053-nathalie-rousseau.html

Carnet de recherche : <https://etymologika.hypotheses.org/>

À votre tour !

Présentation du Consortium Huma-Num ARIANE

Quels enjeux ?

« L'objectif du consortium ARIANE est de progresser dans la connaissance et le raffinement des méthodes informatiques appliquées aux objets et données des sciences humaines et plus particulièrement, des **sciences du texte** ; il veut constituer un espace de discussion portant sur l'interprétation des résultats obtenus à l'aide de méthodes (semi)automatiques d'analyse des textes, car le sens et la validité des extractions effectuées ne sont pas immédiats. »

→ Lien vers la synthèse du projet scientifique du Consortium ARIANE :
<https://shs.hal.science/halshs-04060828>
(voir courriel)

Présentation du Consortium Huma-Num ARIANE

Les groupes de travail ont pour objectif « de tenter de répondre à une question ou à un problème théorique, conceptuel, méthodologique ou technologique ».

Quelles exigences ?

→ Une **participation active des membres**,
une contribution à au moins **un livrable par an**.

N.B. **livrable pour 2023** : une liste d'objectifs scientifiques et un calendrier prévisionnel des activités à rendre le **15 septembre 2023**

Présentation du GT et de son organisation

Plusieurs types d'activités...

1. Réflexions et échanges de pratiques.
 2. **Un recensement critique des outils et méthodes** dédiés à la production d'éditions numériques de qualité (bibliographie, logiciels, méthodes, ressources et données)
 - Mise en place d'un **groupe Zotero partagé**
 - Élaboration de **tutoriels ou de scripts**
 3. **Co-rédaction d'articles** de synthèse ou de réflexion méthodologique
 4. **L'organisation de formations ou d'ateliers méthodologiques**
(sans restrictions de publics ou d'outils)
- + Diffusion (Livrables possibles indiqués en gras)

Remarque : il est possible d'établir des **partenariats**.

Nos premiers contacts sont l'IDHN et le réseau METALEX, d'autres suivront.

Présentation du GT et de son organisation

... sur plusieurs types de sujets différents

Propositions d'ateliers thématiques et transverses

- 0) problématiques transverses
- 1) éditions de textes antiques
- 2) éditions de documents lexicographiques
- 3) autres : vos idées, à préciser ensemble

Ateliers thématiques / transverses

0) Problématiques transverses

Exemples

Comment et pourquoi publier sur ? TEI Publisher, Base X, TXM, mais aussi OMEKA-S, Verbatim...

Comment effectuer un encodage semi-automatique des documents ?

→ L'idée étant de faire état d'une pluralité d'outils, pour une pluralité de besoins.

Ateliers thématiques / transverses

1) Éditions de textes antiques

problématiques qui peuvent être communes à d'autres types de textes anciens:

testimonia, apparats critiques, pas de texte original (?),
multiples éditions et traductions, lemmatisation

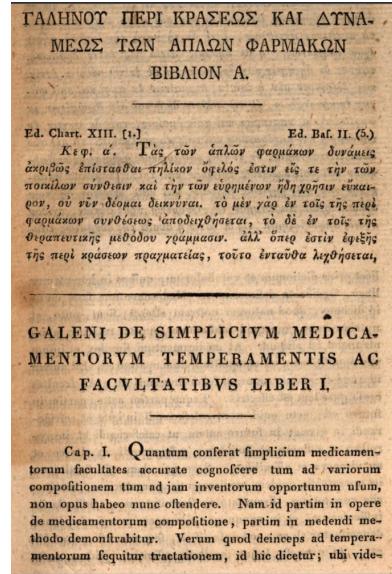
problématiques propres aux textes antiques:

Données:

- « TLG » (Thesaurus Linguae Graecae): tout le corpus de grec ancien, mais fermé, impossible de copier, encore moins de formuler une requête
 - Perseus + « First Thousand Years of Greek » (Scaife)

TEI: normes EpiDoc

citation des textes anciens: The Canonical Text Services URN specification (ex.: urn:cts:greekLit:tlg0057.tlg075.1st1K-grc1:1.1)



<div type="edition" xml:lang="grc" n="urn:cts:greekLit:tlg0057.tlg075.1st1K-grc1">
<pb n="11.379"/>
<div type="textpart" subtype="book" n="1">
<div type="textpart" subtype="chapter" n="1">
<head>ΤΑΝΝΟΥ ΠΕΡΙ ΚΡΑΣΕΩΣ ΚΑΙ ΔΥΝΑΜΕΩΣ
lb/>ΤΟΝ ΑΙΓΑΛΟΝ ΦΑΡΜΑΚΟΝ
lb/>ΒΙΒΛΙΟΝ Α.
</head>
<p>
lb n="1"/> Τάς τῶν ἀπλῶν φαρμάκων δυνάμεις
lb n="2"/> ἀρχιβάς ἐπίστασθαι πηλίκων ὄφελός ἔστιν εἰς τε τὴν τῶν
lb n="3"/> ποικίλων συνθέσιν καὶ τὴν τῶν εὐηρμένων ήδη χρῆσιν εὐάκιρον,
lb n="4"/> οὐν νῦν δέομαι δεικνύναι, τὸ μὲν γάρ ἐν τοῖς τῆς περὶ¹
lb n="5"/> φαρμάκων συνθέσεων ἀποδειχθέσται, τὸ δὲ ἐν τοῖς τῆς
lb n="6"/> θεραπευτικῆς μεθόδων γράμμασιν. ἀλλ' ὅπερ εστὶν ἐφεξῆς
lb n="7"/> τῆς περὶ κράσεων πραγματείας, τούτῳ ἐνταῦθα λεχθήσεται,
lb n="11.380"/>
<div type="edition" xml:lang="grc" n="urn:cts:greekLit:tlg0057.tlg075.st1K-lat1">
<pb n="11.379"/>
<div type="textpart" subtype="book" n="1">
<div type="textpart" subtype="chapter" n="1">
<head>GALENI DE SIMPLICIVM MEDICAMENTORVM
lb/>TEMPERAMENTIS AC
lb/>FACVLTATIBVS LIBER I.</head>
<p>
lb n="1"/> Quantum conserat simplicium medicamenfum
lb n="2"/> facultates accurate cognoscere tum ad variorum
lb n="3"/> compositionem tum ad jam inventorum oportunum usum,
lb n="4"/> non opus habeo nunc ostendere. Nam id partim in opere
lb n="5"/> de medicamentorum compositione , partim in medendi methodo
lb n="6"/> demonstrabitur. Verum quod deinceps ad temperamentorum
lb n="7"/> sequitur tractationem, id hic dicetur; ubi villos
lb n="11.380"/>

<https://galenus-verbatim.huma-num.fr/>

Ateliers thématiques / transverses

2) Éditions de documents lexicographiques

Pour **différents types de documents** :

- Dictionnaires de tous types (monolingues, bilingues, de langue, de spécialité, etc.)
- Encyclopédies
- Lexiques, glossaires, index ou encore annuaires

1) **Réflexions théoriques** d'ensemble :

- Lexicographie, métalexicographie à l'ère du numérique.
- Conception et usage des ouvrages nativement numériques.

2) **Réflexions pratiques** autour de l'encodage :

- Développement d'une chaîne/outil/application permettant la génération semi-automatisée d'un document lexicographique balisé.
- Proposer un enrichissement des standards existants afin de prendre en compte les spécificités des ouvrages anciens (ex. renvois orphelins, diversité des marquages) et appréhender la diachronie (ex. révisions successives, innovations lexicales).



Ateliers thématiques / transverses

3) Vos idées !

Calendrier prévisionnel

- 1. Formation d'initiation à la TEI (standards et encodage, partage des données)** pour les doctorants d'IDF au sein de l'école d'hiver organisée par l'IDHN (Institut des Humanités Numériques) de Cergy le **21 novembre 2023**.
- 2. Journée d'étude du GT3** organisée en partenariat avec l'IDHN (**fin 2023 - début 2024**).
→ Communications des intervenants en présentiel, possibilité d'assister à distance
- 3. Atelier portant sur les dictionnaires numériques** au sein du colloque international organisé par le réseau METALEX les **20 et 21 juin 2024**.

N.B.: Assemblée générale du consortium ARIANE le mercredi **8 novembre 2023, 9h-10h30** (adossée à une formation du GT 4) [**à Lyon, mais accessible en visioconférence**]

Informations pratiques

- Le financement est limité mais possible pour certains déplacements (il est à voir au cas par cas avec les coordinatrices du Consortium).
- Les membres actifs bénéficient d'un accès prioritaire aux formations soutenues.
- Pour des raisons administratives, il est nécessaire d'adhérer au Consortium si vous souhaitez participer activement au GT.

[\(https://framaforms.org/demande-dadhesion-au-consortium-hn-ariane-1679997583\)](https://framaforms.org/demande-dadhesion-au-consortium-hn-ariane-1679997583)

- Les auditeurs et invités aux travaux du GT peuvent s'inscrire à la liste de diffusion pour se tenir au courant des actualités.

[\(https://groupes.renater.fr/sympa/info/infos-cst-ariane\)](https://groupes.renater.fr/sympa/info/infos-cst-ariane)

Tour de table : vos propositions d'ateliers / de livrables / d'activités pour 2024 ou pour une année ultérieure ?

- Vos **propositions pour 2024 (avec le nom d'un référent)** peuvent être accueillies (par courriel) jusqu'au **vendredi 15 septembre**.
- Nous vous proposerons une (seule) réunion (en distanciel) le **lundi 11 ou mercredi 13 septembre** pour en discuter. Un lien Evento vous sera envoyé début septembre.
- Chaque atelier pourra ensuite organiser ses travaux comme il le souhaite.
- Ce groupe de travail est encore ouvert pour le moment : n'hésitez pas à contacter des collègues qui vous sembleraient intéressés.