



GATES
Métiers de
l'ingénierie des
données - retours
d'expérience

GATES

Grenoble Attractiveness and ExcellencES

<https://anr.fr/ProjetIA-22-EXES-0001>
2022-2030

Maxence Larrieu, PhD
GRICAD, UGA, CNRS, INRIA



Objectif



Accroître l'attractivité de la recherche en sciences humaines et sociales, par le prisme des données de recherche

Organisation

Data community manager

Pool
ingénieurs

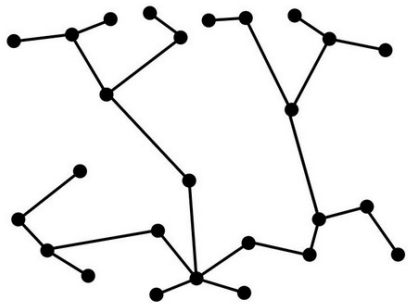


2024, 5 ingénieurs
2025, 7 ingénieurs

Organisation



Projets
de recherche



25 unités

Data community manager



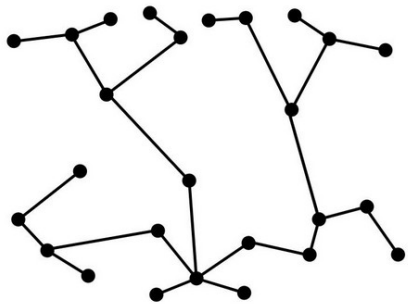
Pool
ingénieurs



Organisation



Projets
de recherche



25 unités



Comité pilotage



Comité scientifique

Data community manager

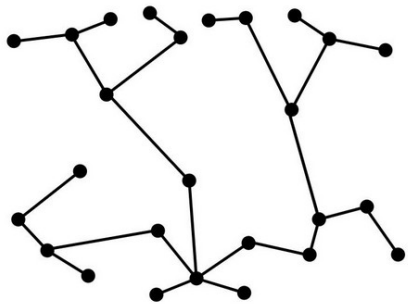
Pool
ingénieurs



Organisation



Projets
de recherche



25 unités



Comité pilotage



Comité scientifique

Data community manager

Pool
ingénieurs



GRICAD

nova,
winter,
summer,
bigfoot,
luke...

DSIM



Translatoscope

Joseph Beau-Reder (GRICAD)
Nassim Mandhouj (Litt&Arts, ARSH)

Projet Translatoscope

Objectif

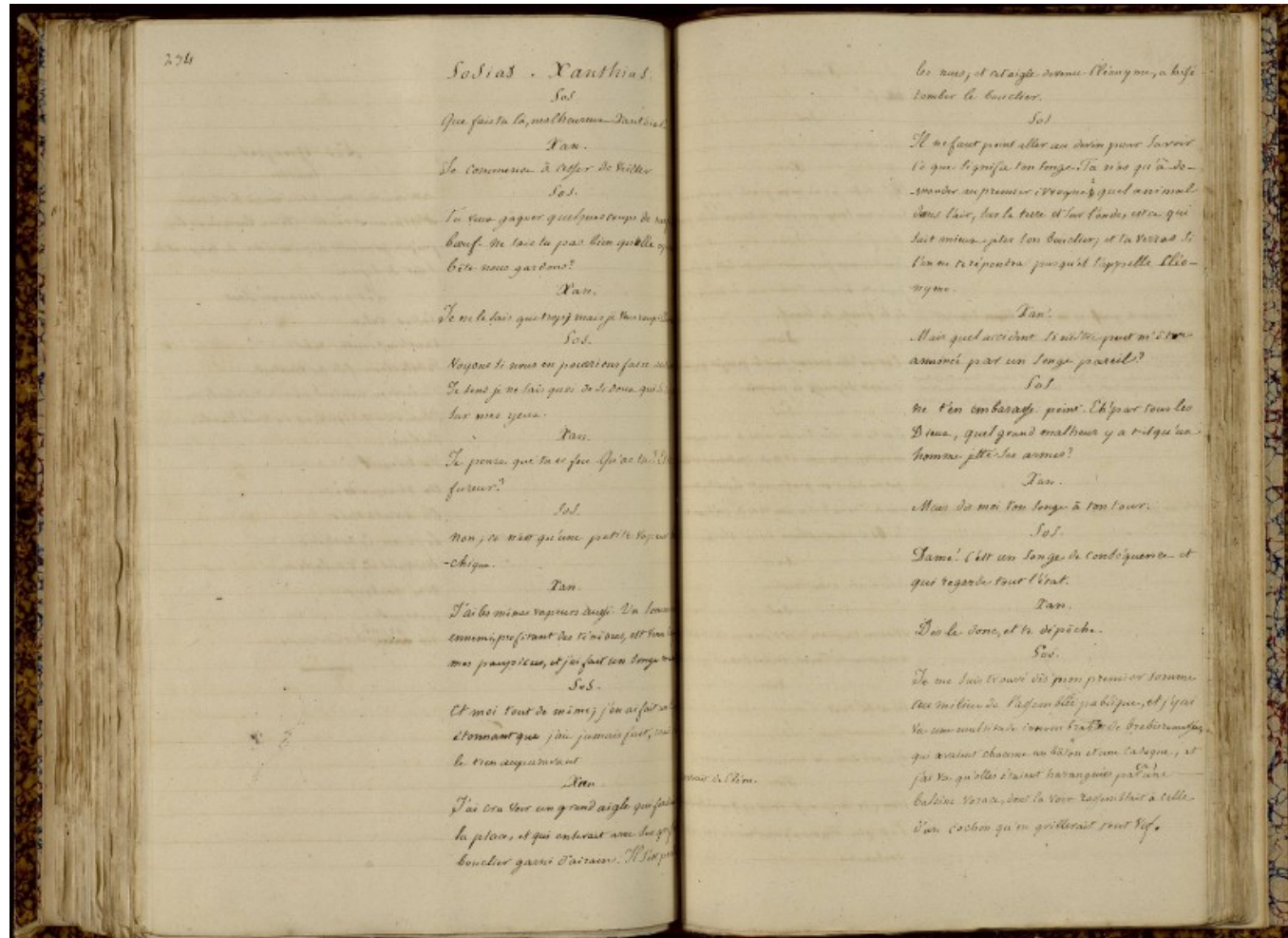
Collecter, numériser, aligner et comparer des traductions de pièces de théâtre grec antique afin d'analyser l'évolution des pratiques de traduction

Contexte scientifique

- ANR *Histoire des traductions en langue française*
- *Translating Greek Drama* (UGA, APGRD, MFO)
- *Les aventuriers de la traduction. Traduire Aristophane en France XVI-XVIIIe s.* Habilitation à diriger des recherches de Malika Bastin-Hammou soutenue en 2015
- *L'assemblée des Femmes* au bac de Littérature, Langues et Culture de l'Antiquité 2024-2026

Origines du projet

L'Aristophane de Lobineau



Origines du projet

L'Aristophane de Lobineau



(a) Le poète raille sur ce qu'Euripide aura dit en quelqu'endroit : Il est inutile de vous dire ce que vous allez voir dans un moment

310

Mnésiloque, Euripide.

Mnésiloque.

O ! Jupiter ! entendra-t-on bientôt gazouiller les Hirondelles ? Cet homme me tue, en me faisant roder dès la pointe du jour. Sera-t-il permis, avant que j'aie perdu toute ma rate à force de courir, de savoir de toi, mon cher Euripide, où tu me mènes ?

Euripide.

Il est inutile que tu entendes ce que tu dois voir tout à l'heure.

Mnési.

Dis-le encore une fois, que je comprenne ^(a). Je ne dois pas entendre

Euri.

Ce que tu dois voir tout à l'heure.

Mnési.

Il ne faudrait donc pas non plus, que je visse.

Euri.

Ce que tu devrais entendre ? Non.

Mnesi.

Que dis-tu là ? Cela est si subtil que je n'y entends pas. Et ce que tu veux, que je ne voie ni n'entende à cause que l'un et l'autre sont incompatibles?

Euri.

Cela est vrai.

Mnésiloque.

Μνησίλοχος

ὦ Ζεῦ χελιδῶν ἄρα ποτε φανήσεται;
ἀπολεῖ μ' ἀλοῶν ἄνθρωπος ἐξ ἐωθινοῦ.
οἶόν τε, πρὶν τὸν σπλῆνα κομιδῆ μ' ἐκβαλεῖν,
παρὰ σοῦ πυφέσθαι ποῖ μ' ἄγεις ὠρπιτίδης;

Εὐριπίδης

5 ἀλλ' οὐκ ἀκούειν δεῖ σε πάνθ' ὅσ' αὐτίκα
ὄψει παρεστῶς.

Μνησίλοχος

πῶς λέγεις; αὐθις φράσον.
οὐ δεῖ μ' ἀκούειν;

Εὐριπίδης

οὐχ ἄ γ' ἂν μέλλης ὀρᾶν.

Μνησίλοχος

οὐδ' ἄρ' ὀρᾶν δεῖ μ';

Εὐριπίδης

οὐχ ἄ γ' ἂν ἀκούειν δέη.

Μνησίλοχος

10 πῶς μοι παραινεῖς; δεξιῶς μέντοι λέγεις.
οὐ φῆς σὺ χρῆναί μ' οὐτ' ἀκούειν οὔθ' ὀρᾶν;

Εὐριπίδης

χωρὶς γὰρ αὐτοῖν ἑκατέρου 'στὶν ἡ φύσις.

Μνησίλοχος

τοῦ μήτ' ἀκούειν μήθ' ὀρᾶν;

Εὐριπίδης

οὐδ' ἄρ' ὀρᾶν;

HAUT

Origines du projet

Traduire Aristophane en Français (TAF)

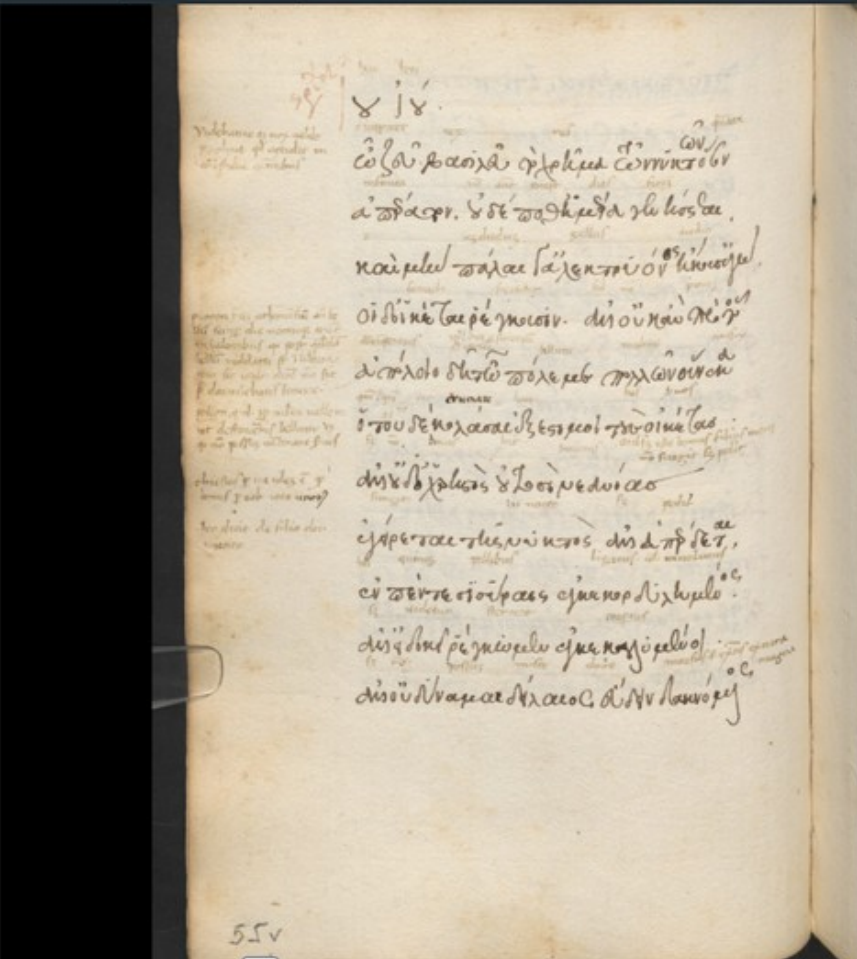
4	XP.	1	(4→)	CHREMYLE
	²⁹ 'Ἄλλ' οὐ σε κρύψω· τῶν ἐμῶν γὰρ οἰκτεῶν πιστότατον ἠγοῦμαι σε καὶ κλεπτίστατον. Ἐγὼ θεοσεβῆς καὶ δίκαιος ὢν ἀνὴρ κακῶς ἔπραττον καὶ πένης ἦν.		Hé bien, je vais te le dire. Car de tous mes gens tu es à mon avis le plus fidelle. ³⁰	
			A part.	
			Je veux dire le plus filou. Sçache donc que tant que j'ai été juste et craignant les Dieux, j'ai vécu gueux et miserable. ¹	
5	KA.	2		CARION
	³⁰ Οἶδά τοι.		Oh, je sçai cela.	
6	XP.	3		CHREMYLE
	³¹ Ἄτεροι δ' ἐπλούτουν. ἱερόσυλοι, ῥήτορες καὶ συκοφάνται καὶ πονηροί.		Pour les autres, par exemple, sacrileges, orateurs, délateurs, scélérats de toute espece, je les ai vûs riches.	
7	KA.	4		CARION
	Πείθομαι.		Je le croi bien vraiment.	
8	XP.	5		CHREMYLE
	³² Ἐπερησόμενος οὖν ᾠχόμην ὡς τὸν θεόν, τὸν ἐμὸν μὲν αὐτοῦ τοῦ ταλαιπώρου σχεδὸν ἤδη νομίζων ἐκτετοξεῦσθαι βίον, τὸν δ' υἱόν, ὅσπερ ὢν μόνος μοι τυγχάνει, πευσόμενος εἰ χρὴ μεταβαλόντα τούς τρόπους εἶναι πανούργου, ἀδικου, ὑγιές μηδὲ ἐν, ὡς τῷ βίῳ τοῦτ' αὐτὸ νομίσας ξυμφέρειν.		Je me suis donc avisé d'aller consulter l'oracle, comme étant sur la fin de mes vieux jours et de ma misere, pour sçavoir si le fils unique que j'ai ne feroit pas mieux de changer de train, pour devenir fourbe, injuste, et mechant; puisque c'est le vrai moien d'être heureux.	
9	KA.	6		CARION
	⁴⁰ "Τί δῆτα Φοῖβος ἔλακεν ἐκ τῶν στεμμάτων;"		Hé, qu'a répondu le Dieu du fonds de ses épais lauriers ?	
10	XP.	7		CHREMYLE
	Πεύσει. Σαφῶς γὰρ ὁ θεὸς εἶπέ μοι ταδί· ὄψ ξυναντήσαμι πρῶτον ἐξῶν, ἐκέλευε τούτου μὴ μεθίσθαι μ' ἔτι, πείθειν δ' ἐμαυτῷ ξυνακολουθεῖν οἴκαδε.		Il m'a dit de m'attacher au premier homme que je trouverois à l'issue du Temple, de ne le pas quitter, et de l'engager à me suivre chés moi.	
11	KA.	8		CARION

Renaissance Aristophanes Digital (RenArD)

L'Aristophanes Latinus d'Alexandre d'Otrante, Nubes

55v

BUTTONS.MSD



55v

55v

Doc 1

55v

Diplomatic

BUTTONS.INFO ⓘ

STREPSIADES:

Heu, heu!

O Iuppiter rex, res _ _ _ _ quanta:
infinita. Non autem potest dies fieri?

Et iam dudum gallus audio.

Famuli stertunt. Sed non primum.

Corrumpereris igitur, dite per sincopam, aliter utinam, bellum, multorum causa,
quoniam, vel quando, non autem cruciari licet _ _ _ _ hos servos.

Sed non bonus • hic iuvenis

surgit in nocte, sed pedit,

in quinque pellibus ligatus, vel involutus.

Sed _ _ _ _ videtur, stertere • coopertus.

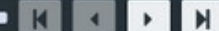
Sed non possum, miser, dormire morsus, scilicet existens ex cura magna,

Selects.no SELECTION

Search

Create index

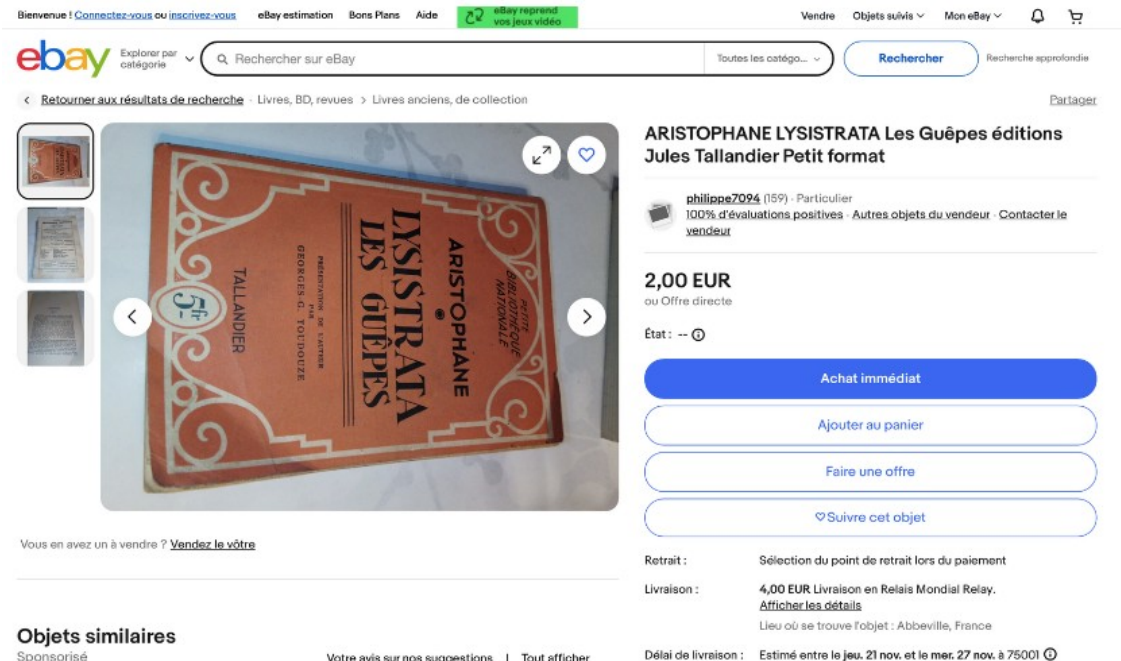
A⁺



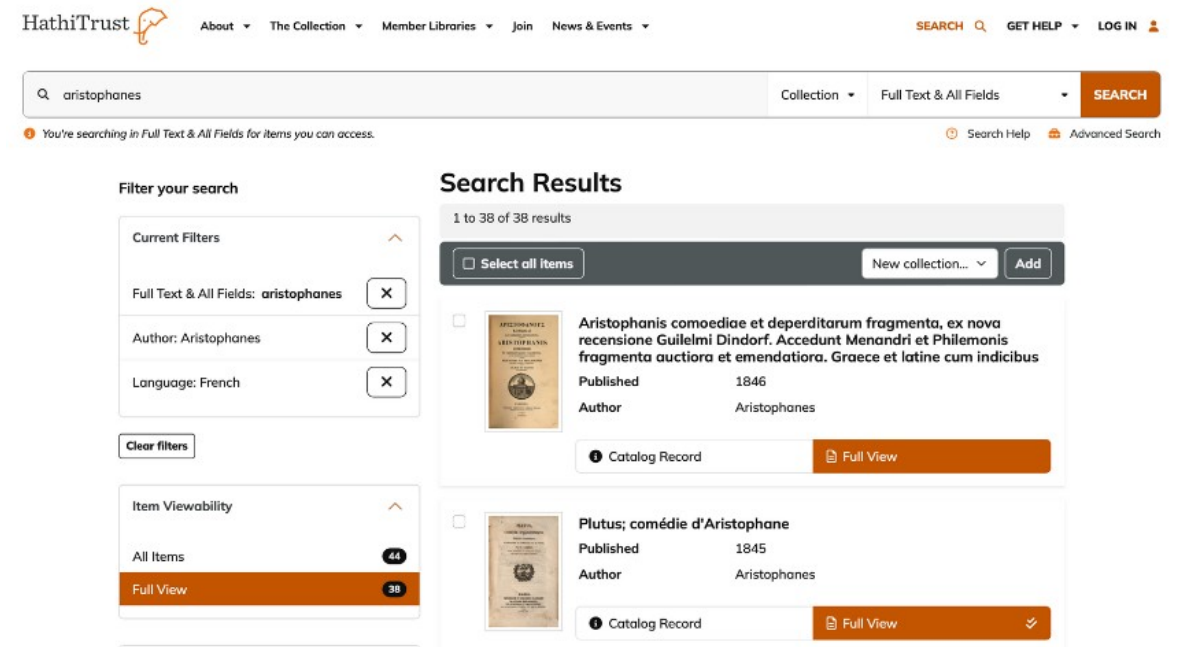
Produire une édition numérique

étape 1 : Collecte bibliographique

Recherche de mots clés, comme “Aristophane” ou le nom des pièces avec les variantes orthographiques liées aux pratiques de chaque époque ex : Guêpes - Guespes, sur les bases de données telles que “Google books”, “Gallica”, “HathiTrust” ou même e-market tel que “ebay”.



The screenshot shows an eBay product page for a book. The main image is a decorative book cover with the title 'ARISTOPHANE LYSISTRATA LES GUÊPES' and the publisher 'TALLANDIER'. The price is listed as 2,00 EUR. The seller is 'philippe7094' with a 100% positive feedback rating. The page includes navigation buttons like 'Achat immédiat', 'Ajouter au panier', and 'Faire une offre'. At the bottom, there are sections for 'Objets similaires' and 'Sponsorisé'.



The screenshot shows the HathiTrust search results for the query 'aristophanes'. The search filters are set to 'Full Text & All Fields'. The results show two items: 'Aristophanis comoediae et deperditarum fragmenta, ex nova recensione Guilelmi Dindorf. Accedunt Menandri et Philemonis fragmenta auctiora et emendatiora. Graece et latine cum indicibus' (published 1846) and 'Plutus; comédie d'Aristophane' (published 1845). Both items have 'Full View' buttons available.

Produire une édition numérique

étape 1 : Collecte bibliographique

Recherche de mots clés, comme “Aristophane” ou le nom des pièces avec les variantes orthographiques liées aux pratiques de chaque époque ex : Guêpes - Guespes, sur les bases de données telles que “Google books”, “Gallica”, “HathiTrust” ou même e-market tel que “ebay”.

The screenshot shows a Google search for "les guêpes" Aristophane. The search results are displayed in a dark theme. The top result is "Les guêpes" by Aristophane, published in 2012. The description states: "Les guêpes" sont notre sixième publication des œuvres complètes d'Aristophane. Below the title, there are buttons for "Aperçu" and "Autres éditions". The second result is "Aristophane: Les guêpes. La paix" by Aristophanes, published in 1962. It is noted as an "Affichage d'extraits" and has an "Autres éditions" button. The third result is "Plutus. Les guêpes. Les fêtes de Cérés" by Aristophanes, published in 1790. It has "Lire" and "Autres éditions" buttons.

The screenshot shows the Gallica website search results for "Aristophane". The search bar contains "Rechercher..." and the results page shows 85 results on page 1. The search filters on the left include "Ma recherche" (Auteur/Contributeur: Aristophane), "RESULTATS" (85 documents consultables en ligne, 63 sur place), and "Affiner" (Lancer la recherche dans ces résu). The "Site de consultation" list includes Gallica (59), e-rara.ch (11), numelyo (5), Bibliothèques virtuelles humanistes (2), e-codices (2), and Bibliothèque interuniversitaire de la Sorbon... (2). The "Type de document" list includes Livres (45), Manuscrits (37), and Images (3). The main results list shows three items:

3. **Plutus, ou La richesse / comédie d'Aristophane ; traduit du grec en vers français, par Eugène Fallex,...** 1849
Auteur **Aristophane** (04457-03867 av. J.-C.). Auteur du Texte
Informations détaillées
4. **Plutus, ou L'égalité répartition des richesses : comédie en vers / traduite d'Aristophane, par Amédée Fleury** 1851
Auteur **Aristophane** (04457-03867 av. J.-C.). Auteur du Texte
Informations détaillées
5. **Les oiseaux : fantaisie en 2 actes d'après Aristophane / Fernand Nozière ; [traduction Lascaris]** 1911
Contributeur **Aristophane** (04457-03867 av. J.-C.). Auteur adapté
Informations détaillées

The sixth result is partially visible: "Les comices d'Athènes, ou Les femmes orateurs, comédie vaudeville en un acte, traduit du grec (d'Aristophane), par MM. Eugène Scribe et ***" 1817.

Produire une édition numérique

étape 2 : Automatic text recognition (ATR)

Grâce à l'outil d'ATR eScriptorium nous avons pu récupérer les textes contenus dans les fichiers pdf des traductions pour les transformer en format text.

The screenshot displays the eScriptorium web application interface. At the top, there is a navigation bar with 'Home' and 'Contact' links, a search bar containing 'Search in brotier guêpes', and user information 'My Projects My Models Hello Nassim'. Below this, a breadcrumb trail shows 'Description Ontology Images Edit Models Reports' for the document 'brotier guêpes'. A search bar contains the text 'kraken:Gallicorpora+_best'. The main content area is divided into three columns. The left column shows the original PDF text with a search icon. The middle column shows the same text with red bounding boxes around the text, indicating the ATR process. The right column shows the extracted text in a clean, digital format. The text in the middle column is highlighted in red, and the text in the right column is in a clean, digital format.

Produire une édition numérique

étape 2 : Automatic text recognition (ATR)

1) Implanter le texte dans l'interface et l'analyser afin de choisir le meilleur model d'OCR disponible.

2) Segmenter, càd délimiter les lignes et les zones que l'on souhaite transcrire.

3) Transcription automatique. Ici à l'aide du modèle "Gallicorpora+_best".

The screenshot shows the eScriptorium web interface. At the top, there is a navigation bar with 'Home' and 'Contact' links, a search bar containing 'Search in brotier guêpes', and user information 'My Projects My Models Hello Nassim'. Below this, a breadcrumb trail shows 'Description Ontology Images Edit Models Reports' for the document 'brotier guêpes'. The main content area is divided into three columns. The left column shows the original scanned text from a book, titled 'LES GUËPES. 283'. The middle column shows the same text with blue boxes around each line, indicating segmentation. The right column shows the original text again, but with some lines highlighted in blue, indicating the transcribed text. The interface includes various navigation and editing tools, such as a search icon, a download icon, and a toolbar with icons for zooming, panning, and other document manipulation functions.

Produire une édition numérique

étape 3 : Relecture des sorties d'OCR

Enregistrement automatique E_ Enregistré dans mon Mac

Accueil Insertion Dessin >> Commentaires Modification Partager >>

Coller Police Paragraphe Styles Dictier Sensibilité Compléments Rédacteur

UN JEUNE HOMME.
UNE VIEILLE.
MERCURE.
LE PRÊTRE de Jupiter.
QUELQUES PERSONNAGES MUETS.

La scène est à Athènes, devant la maison de Chrémyle.

PLUTUS'.
COMÉDIE D'ARISTOPHANE.
ACTE PREMIER.
SCÈNE PREMIÈRE.
CARION, CHRÉMÉYLE, PLUTUS.

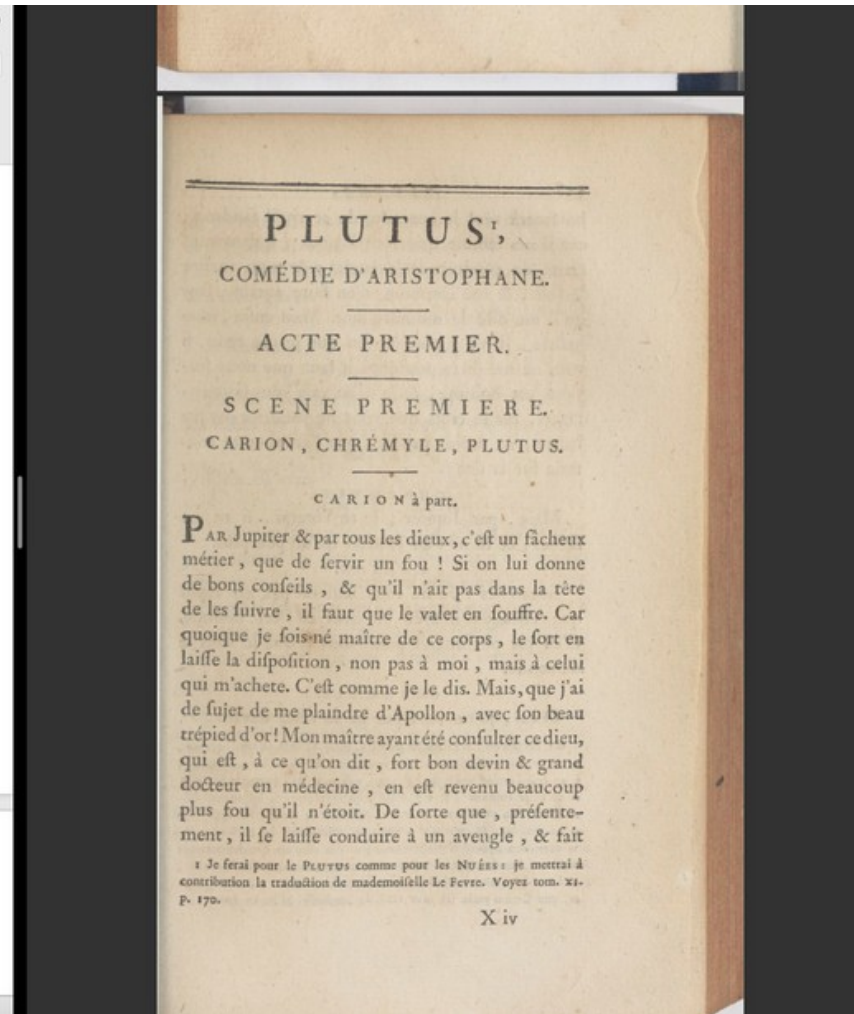
CARION à part.
PAR Jupiter et par tous les dieux, c'est un fâcheux métier, que de servir un fou ! Si on lui donne de bons conseils, et qu'il n'ait pas dans la tête de les suivre, il faut que le valet en souffre. Car quoique je sois né maître de ce corps, le sort en laisse la disposition, non pas à moi, mais à celui qui m'achète. C'est comme je le dis. Mais, que j'ai de sujet de me plaindre d'Apollon, avec son beau trépied d'or ! Mon maître ayant été consulter ce dieu, qui est, à ce qu'on dit, fort bon devin et grand docteur en médecine, en est revenu beaucoup plus fou qu'il n'étoit. De sorte que, présentement, il se laisse conduire à un aveugle, et fait justement tout le contraire de ce qu'il faudroit ; car il me semble que c'est à nous, qui voyons clair, de conduire les aveugles : et mon maître le sait, et me contraint d'en faire autant, sans qu'il me dise le moindre mot. Mais enfin, mon maître, il n'y a plus

¹ Je ferai pour le PLUTUS comme pour les NUÉES : je mettrai à contribution la traduction de mademoiselle Le Févre. Voyez tom. XI. P. 170.

moyen que je me taise, si vous ne me dites pourquoi il faut que nous suivions cet homme, et je m'en vais vous faire enrager, car je crois que vous ne voudriez pas me battre, au moins pendant que j'aurai cette couronne sur la tête.

CHRÉMÉYLE.

Page 1 sur 53 17221 mots Focus 125 %



Produire une édition numérique

étape 4 : structuration en XML-TEI

Lien vers l'ODD



Header contenant les métadonnées.



```
processing-instruction
1 <?xml version="1.0" encoding="UTF-8"?>
2 <?xml-model href="https://gricad-gitlab.univ-grenoble-alpes.fr/elan/schemas-tei/-/raw/main/translatoscope_taf/out/taf_odd.mg?ref_type=heads" type="application/xml" schematypens="http://relaxng.org/ns/structure/1.0"?>
3 <?xml-model href="https://gricad-gitlab.univ-grenoble-alpes.fr/elan/schemas-tei/-/raw/main/translatoscope_taf/out/taf_odd.mg?ref_type=heads" type="application/xml" schematypens="http://purl.oclc.org/dtd/schematron"?>
4 <TEI xmlns="http://www.tei-c.org/ns/1.0">
5   <teiHeader>
6     <fileDesc>
7       <titleStmt>
8         <title type="file">PLUTUS : une édition numérique du texte grec traduit en français
9           par André Charles Brotier</title>
10        <respStmt>
11          <resp>Responsable du projet</resp>
12          <name>Malika Bastin</name>
13        </respStmt>
14        <respStmt>
15          <resp>Nettoyage des sorties d'OCR</resp>
16          <name>Nassim Mandhouj</name>
17        </respStmt>
18        <respStmt>
19          <resp>Transformation en XML-TEI</resp>
20          <name>Nassim Mandhouj</name>
21        </respStmt>
22      </titleStmt>
23      <publicationStmt>
24        <publisher>Littéart/Arts</publisher>
25        <pubPlace>Université Grenoble Alpes</pubPlace>
26      </publicationStmt>
27      <sourceDesc>
28        <pb>
29          <bibl>
30            <title type="play">Plutus</title>
31            <author>Aristophane</author>
32            <editor role="traducteur">André Charles Brotier</editor>
33            <title type="main">Théâtre des Grecs, par le P. Brunoy. Nouvelle édition,
34              enrichie de très-belles gravures, et augmentée de la traduction entière
35              des pièces grecques, dont il n'existe que des extraits dans toutes les
36              éditions précédentes, et de comparaisons, d'observations et de remarques
37              nouvelles, par MM. de Rochefort et Du Theil, ... et par M***</title>
38            <date>1789</date>
39            <pubPlace>Paris</pubPlace>
40            <publisher>Cuseac</publisher>
41            <biblScope unit="volume">13</biblScope>
42          </bibl>
43        </pb>
44      </sourceDesc>
45    </fileDesc>
46    <profileDesc/>
47  </teiHeader>
48  <text>
49    <front>
50      <div type="didascalie">
51        <pb/>
52        <head>PERSONNAGES</head>
53      </div>
54      <castlist>
55        <castitem corresp="XERXES">CHÉRÉMÈLE.</castitem>
56        <castitem corresp="MÉLITE">CÉCILE, valet de CHÉRÉMÈLE.</castitem>
57        <castitem corresp="MÉLITE">PLUTUS.</castitem>
58        <castitem corresp="MÉLITE">OHIBR DE VILLAGEOIS.</castitem>
59        <castitem corresp="MÉLITE">BLEPSISDÈME, ami de CHÉRÉMÈLE.</castitem>
60      </castlist>
61    </front>
62  </text>
63 </TEI>
```


Produire une édition numérique

étape 4 : structuration en XML-TEI

Toute la traduction est contenue dans une balise <body>



Les balises <div> permettent de délimiter les différentes parties



Chaque réplique est structurée de la même manière c`ad : une réplique est contenue dans une balise <sp>, la personne qui parle dans une balise <speaker> et le texte dans une balise <p>



```
TEI text body div head
75 <body>
76   <div n="PLUTUS" subtype="edition_text" type="drama" xml:id="PLUTUS">
77
78     <head type="play">PLUTUS,<lb/> COMÉDIE<lb/> D'ARISTOPHANE.</head>
79     <lb/>
80     <div type="partie" n="1">
81       <head>ACTE PREMIER.</head>
82       <div type="scene" n="1">
83         <head>SCÈNE PREMIÈRE</head>
84         <stage>CARION, CHRÉMÉYLE, PLUTUS.</stage>
85
86
87         <sp>
88           <speaker corresp="KAPION">CARION.</speaker>
89           <stage>à part.</stage>
90
91           <p>PAR Jupiter et par tous les dieux, c'est un fâcheux métier, que de
92             servir un fou ! Si on lui donne de bons conseils, et qu'il n'ait pas
93             dans la tête de les suivre, il faut que le valet en souffre. Car
94             quoique je sois né maître de ce corps, le sort en laisse la
95             disposition, non pas à moi, mais à celui qui m'achete. C'est comme
96             je le dis. Mais, que j'ai de sujet de me plaindre d'Apollon, avec
97             son beau trépied d'or ! Mon maître ayant été consulter ce dieu, qui
98             est, à ce qu'on dit, fort bon devin et grand docteur en médecine, en
99             est revenu beaucoup plus fou qu'il n'étoit. De sorte que,
100            présentement, il se laisse conduire à un aveugle, et fait justement
101            tout le contraire de ce qu'il faudroit ; car il me semble que c'est
102            à nous, qui voyons clair, de conduire les aveugles : et mon maître
103            le suit, et me contraint d'en faire autant, sans qu'il me dise le
104            moindre mot. Mais enfin, mon maître, il n'y a plus moyen que je me
105            taise, si vous ne me dites pourquoi il faut que nous suivions cet
106            homme, et je m'en vais vous faire enrager, car je crois que vous ne
107            voudriez pas me battre, au moins pendant que j'aurai cette couronne
108            sur la tête.</p>
109
110         </sp>
```

Produire une édition numérique

One Document does it all (ODD)

Offre la possibilité de n'autoriser que certaines balises et éléments. Cela permet d'avoir des fichiers encodés de manière uniforme. Utiliser un ODD est fort utile dans un projet nécessitant des données en XLM TEI et pour lequel différents acteurs interviennent.

Produire une édition numérique

One Document does it all (ODD)

Offre également la possibilité de produire automatiquement un manuel dans lequel chaque balise et élément utilisés peuvent être détaillés et expliqués ce qui favorise l'interopérabilité et la réutilisations des données avec une justification des choix, ici d'édition numérique.

Manuel d'encodage pour la transcription du projet TAF

Table de matières

Élément `TEI`

L'élément racine du fichier XML; indique la norme selon laquelle les données sont encodées.

Élément `teiHeader`

Le `<teiHeader>` contient une série de métadonnées concernant le fichier, ses sources et son encodage.

Élément `fileDesc`

Contient une description bibliographique complète du fichier électronique.

Élément `titleStmt`

Permet d'encoder les différentes informations liées à la fois au fichier numérique et à la traduction latine objet de l'édition.

Élément `title`

Utilisé à l'intérieur de `<titleStmt>`, l'élément `<title>` permet d'indiquer à la fois le nom du fichier numérique et celui de la comédie d'Aristophane traduite en latin.

Attribut :

- `@type`
 - `file` : La valeur "file" définit le titre du fichier numérique.
 - `main`
 - `play` : La valeur "play" indique que le titre est celui de l'œuvre originale.

Élément `respStmt`

Fournit une déclaration de responsabilité pour sur les individus ou les organisations ayant participé à la production et révision de l'édition numérique.

Élément `resp`

Permet d'indiquer les différents responsables ayant collaboré au projet.

Attribut :

- `@when` Indique la date selon le format `aaaa-mm` (ex. 2023-07) à laquelle la modification a été effectuée.

Projet Translatoscope

Objectif

Collecter, numériser, aligner et comparer des traductions de pièces de théâtre grec antique afin d'analyser l'évolution des pratiques de traduction

Soutien apporté par GATES

- créer un outil accessible pour **aligner** les traductions d'une même pièce de théâtre
- créer un site web pour **exposer** les données et **comparer** facilement les traductions
- **ouvrir** les données sur Nakala

La problématique de l'alignement

Traduction A

Carion :
Je vois.

Chrémyle:
C'était simple.

Traduction B

Carion :
Je comprends.

Ploutos :
Enfin...

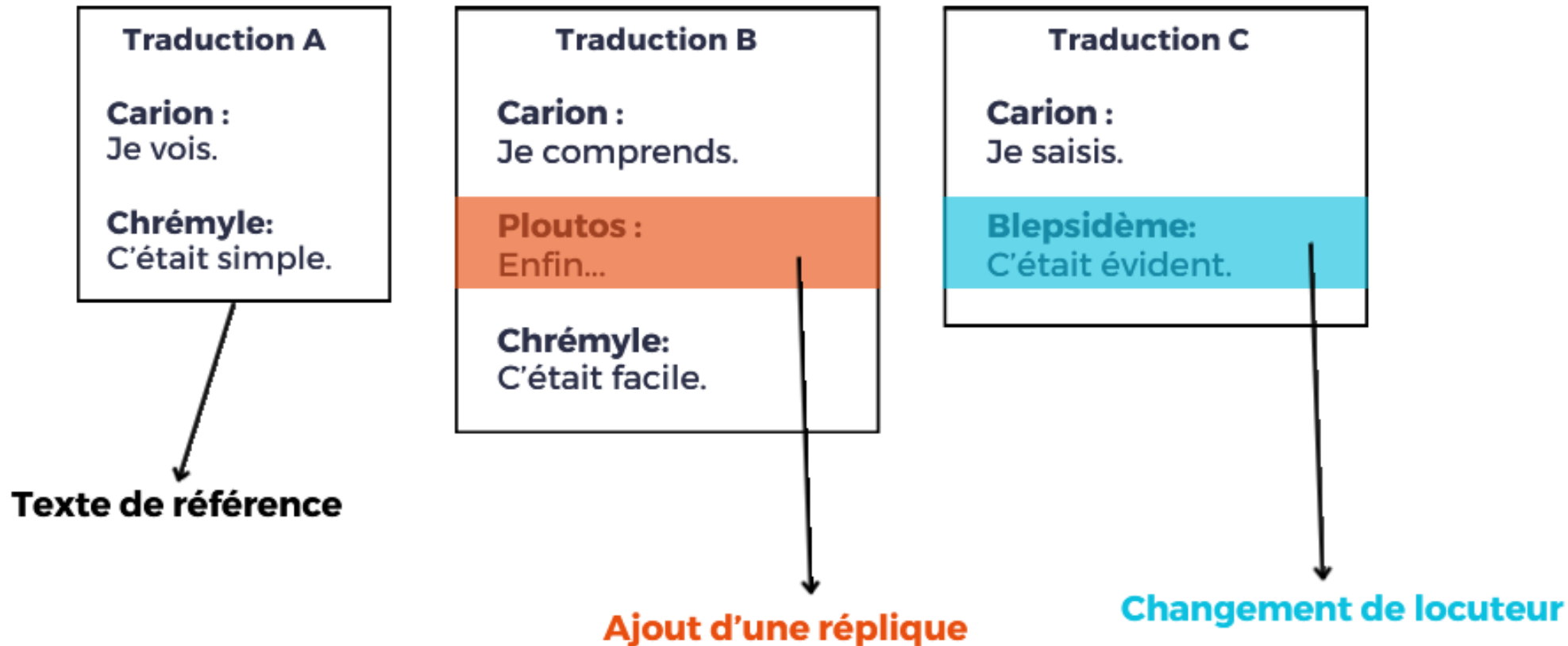
Chrémyle:
C'était facile.

Traduction C

Carion :
Je saisis.

Blepsidème:
C'était évident.

La problématique de l'alignement

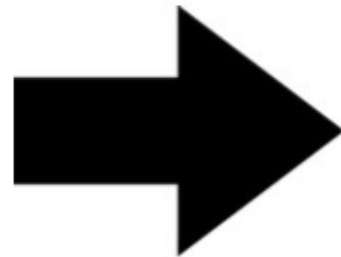


Développer un "outil d'alignement"

Ensemble de codes (python) téléchargeable depuis Gricad-Gitlab



1. Transformer un fichier XML-TEI (.xml) en un tableur LibreOffice/Excel (.xlsx)



Développer un "outil d'alignement"

Ensemble de codes (python) téléchargeable depuis Gricad-Gitlab



1. Transformer un fichier XML-TEI (.xml) en un tableur LibreOffice/Excel (.xlsx)

2. Aligner un texte à l'aide du logiciel LibreOffice Calc



A	B	C	D
1	Personnage Coulon	Personnage VanDaele	Texte VanDaele
2	KARION	CARION	Quelle pénible chose c'est, à Zeus et tous les dieux, de devenir esclave d'un maître hors de son bon sens ! Se trouve-t-il que le serviteur ait donné les meilleurs conseils, et qu'il ait plu à son possesseur de ne pas les<fw>5</fw> suivre, fatalement le serviteur aura sa part des maux..
3	XREMYLOS	CHRÉMYLE	Non, par Zeus ! mais je t'ôterai ta couronne, pour peu que tu m'ennuies, qu'il t'en cuira d'autant plus.
4	KA.	BLEPSIDÈME	Chansons ! Je ne cesserai pas, que tu ne m'aies dit qui peut bien être cethomme. C'est parce que je <fw>25</fw> te veux du bien que je te le demande, le plus grand bien.
5	XP.	CHRÉMYLE	Eh bien, je ne te le cacherai pas, car de nos serviteurs je te tiens pour le plus fidèle et le plus voleur. Moi, quoique religieux et homme juste, je n'éréussissais pas et j'étais pauvre.
6	KA.	CARION	Je le sais, ma foi.

Le site web

elan.gricad-pages.univ-grenoble-alpes.fr/
translatoscope/

Translatoscope [Alignements](#) [Source](#) [A propos](#)

UGA
Université
Grenoble Alpes

Choisir texte ▾ Code couleur

+ -

Ploutos - 1930 - VanDaele ✕

Voilà, nous te lâchons.

PLOUTOS :
Écoutez donc. Car il me faut bien, à ce que je vois, dire ce que j'étais disposé à tenir secret : je suis Ploutos.

CARION :
O le plus scélérat de tous les hommes ! Tu es Ploutos et tu te taisais !

CHRÉMYLE :
Toi Ploutos, en si piteux état ? O Phoibos Apollon, ô dieux, ô démons, ô Zeus, que dis-tu ? Ploutos, réellement, c'est toi ?

PLOUTOS :
Oui.

CHRÉMYLE :
Ploutos en personne ?

PLOUTOS :

Ploutos - 1512 - Beatus Rhenanus ✕

dicere quae celare eram paratus: ego quidem sum Plutus.

Chremylus :
O, scelestissime virorum omnium, deinde tacuisti Plutus existens? Tu Plutus, sic misere habens positus? O Phebe Apollo et dii et daemones et Iupiter, quid dicis? Ille vere es tu?

:

Plutus :
Certe.

Chremylus :
Ille ipse?

Plutus :
Ipsissimus!

Chremylus :

Ploutos - 1684 - Dacier ✕

Hé bien, nous te laissons.

PLUTUS. :
Écoutez maintenant ; car je voy bien qu'il faut que je vous dise ce que j'avois resolu de vous cacher ; je suis Plutus.

CREMYLE. :
O le plus scelerat de tous les hommes ! Tu serois Plutus, et tu nous l'aurois caché ?

CARION. :
Toi, Plutus ? bâti comme te voila ? Ô Phebus ! Apollon ! ô grands Dieux ! Ô toutes les Divinités essemble ! ô Jupiter ! Quoy tu serois effectivement Plutus ?

PLUTUS. :
Sans doute.

CARION. :
Luy-même.

Ploutos - 1849 - Fallex ✕

PLUTUS :
Eh bien, donc, apprenez ce secret important. Je m'étais bien promis de ne pas vous le dire, Mais je vois qu'à la fin il faut vous en instruire. C'est moi qui suis Plutus.

CHRÉMYLE :
O menteur effronté, N'aurais-tu pas parlé, si tu l'avais été ?

CARION :
Toi Plutus ? Allons donc ! pas dans cette débîne !

CHRÉMYLE :
Par Phébus Apollon, dieux et bonté divine, Jupiter ! quoi, Plutus ? tu prétends que c'est toi ?

PLUTUS :
Sûrement

Documentation

Vers une autonomie du laboratoire

1. Télécharger toutes les ressources nécessaires (codes, logiciels...)
2. Utiliser les fichiers de code (aligner une pièce...)
3. Ajouter du contenu au site web

Ouverture des données

Présentation des sources

Données

Les données au format TEI produites dans le cadre du projet Translatoscope sont partagées dans l'entrepôt de données Nakala : <https://nakala.fr/collection/10.34847/nkl.798c6967>.
Les données sont mises à disposition selon les termes de la Licence Ouverte. Citez l'ensemble des données produites :

Malika Bastin (2024) "Translatoscope", collection Nakala. <https://nakala.fr/collection/10.34847/nkl.798c6967>

Les traductions latines sont des traductions inédites d'Aristophane des XVe et XVIe siècles, elles ont été recueillies dans le cadre du projet Translatoscope, pour parcourir visuellement les manuscrits, rendez-vous sur le site suivant : <https://renard.elan-numerique.fr/>.

Code

L'ensemble des codes réalisés pour traiter, aligner et exposer les données sont partagés sur la forge de l'université : <https://gricad-gitlab.univ-grenoble-alpes.fr/elan/translatoscope>.
Les jupyter notebook sont documentés afin de favoriser leur réutilisation.

 **Ouvrir
la science !**

 **GitLab**

Droit et licences

- Deux sites web : un ouvert à tous, un restreint aux membres du projet

Droit et licences

- Deux sites web : un ouvert à tous, un restreint aux membres du projet
- Licence du code : GNU Affero General Public License v3.0



- Licence des données (Nakala) : Etalab Open License 2.0



LICENCE OUVERTE
OPEN LICENCE

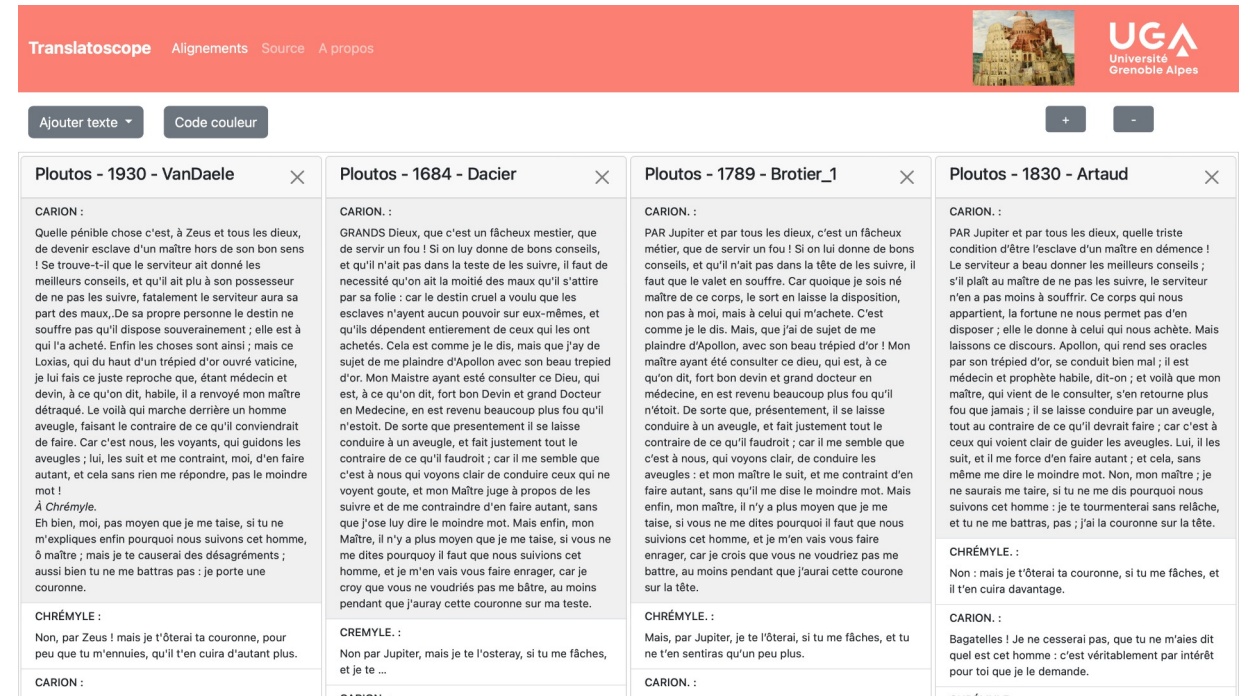
Conclusion : Bilan et perspectives

- GATES et ELAN

- L'outil

- L'interface

- Les données



The screenshot displays the Translatoscope web application interface. At the top, there is a navigation bar with the title "Translatoscope" and links for "Alignements", "Source", and "A propos". The UGA logo is visible in the top right corner. Below the navigation bar, there are two buttons: "Ajouter texte" and "Code couleur". The main content area is divided into four parallel text windows, each with a title and a close button (X):

- Ploutos - 1930 - VanDaele**: CARION : Quelle pénible chose c'est, à Zeus et tous les dieux, de devenir esclave d'un maître hors de son bon sens ! Se trouve-t-il que le serviteur ait donné les meilleurs conseils, et qu'il ait plu à son possesseur de ne pas les suivre, fatalement le serviteur aura sa part des maux...De sa propre personne le destin ne souffre pas qu'il dispose souverainement ; elle est à qui l'a acheté. Enfin les choses sont ainsi ; mais ce Loxias, qui du haut d'un trépid d'or ouvré vaticine, je lui fais ce juste reproche que, étant médecin et devin, à ce qu'on dit, habile, il a renvoyé mon maître détraqué. Le voilà qui marche derrière un homme aveugle, faisant le contraire de ce qu'il conviendrait de faire. Car c'est nous, les voyants, qui guidons les aveugles ; lui, les suit et me contraint, moi, d'en faire autant, et cela sans rien me répondre, pas le moindre mot !
À Chrémyle.
Eh bien, moi, pas moyen que je me taise, si tu ne m'expliques enfin pourquoi nous suivions cet homme, ô maître ; mais je te causerai des désagréments ; aussi bien tu ne me battras pas : je porte une couronne.
- Ploutos - 1684 - Dacier**: CARION : GRANDS Dieux, que c'est un fâcheux mestier, que de servir un fou ! Si on luy donne de bons conseils, et qu'il n'ait pas dans la teste de les suivre, il faut de nécessité qu'on ait la moitié des maux qu'il s'attire par sa folie : car le destin cruel a voulu que les esclaves n'ayent aucun pouvoir sur eux-mêmes, et qu'ils dépendent entièrement de ceux qui les ont achetés. Cela est comme je le dis, mais que j'ay de sujet de me plaindre d'Apollon avec son beau trépid d'or. Mon Maistre ayant esté consulter ce Dieu, qui est, à ce qu'on dit, fort bon Devin et grand Docteur en Medecine, en est revenu beaucoup plus fou qu'il n'estoit. De sorte que presentement il se laisse conduire à un aveugle, et fait justement tout le contraire de ce qu'il faudroit ; car il me semble que c'est à nous qui voyons clair de conduire ceux qui ne voyent goutte, et mon Maître juge à propos de les suivre et de me contraindre d'en faire autant, sans que j'ose luy dire le moindre mot. Mais enfin, mon Maître, il n'y a plus moyen que je me taise, si vous ne dites pourquoi il faut que nous suivions cet homme, et je m'en vais vous faire enrager, car je croy que vous ne voudriez pas me bâtre, au moins pendant que j'auray cette couronne sur ma teste.
- Ploutos - 1789 - Brotier_1**: CARION : PAR Jupiter et par tous les dieux, c'est un fâcheux métier, que de servir un fou ! Si on lui donne de bons conseils, et qu'il n'ait pas dans la tête de les suivre, il faut que le valet en souffre. Car quoique je sois né maître de ce corps, le sort en laisse la disposition, non pas à moi, mais à celui qui m'achete. C'est comme je le dis. Mais, que j'ai de sujet de me plaindre d'Apollon, avec son beau trépid d'or ! Mon maître ayant été consulter ce dieu, qui est, à ce qu'on dit, fort bon devin et grand docteur en médecine, en est revenu beaucoup plus fou qu'il n'étoit. De sorte que, présentement, il se laisse conduire à un aveugle, et fait justement tout le contraire de ce qu'il faudroit ; car il me semble que c'est à nous, qui voyons clair, de conduire les aveugles : et mon maître le suit, et me contraint d'en faire autant, sans qu'il me dise le moindre mot. Mais enfin, mon maître, il n'y a plus moyen que je me taise, si vous ne me dites pourquoi il faut que nous suivions cet homme, et je m'en vais vous faire enrager, car je crois que vous ne voudriez pas me battre, au moins pendant que j'aurai cette couronne sur la tête.
- Ploutos - 1830 - Artaud**: CARION : PAR Jupiter et par tous les dieux, quelle triste condition d'être l'esclave d'un maître en démence ! Le serviteur a beau donner les meilleurs conseils ; s'il plaît au maître de ne pas les suivre, le serviteur n'en a pas moins à souffrir. Ce corps qui nous appartient, la fortune ne nous permet pas d'en disposer ; elle le donne à celui qui nous achète. Mais laissons ce discours. Apollon, qui rend ses oracles par son trépid d'or, se conduit bien mal ; il est médecin et prophète habile, dit-on ; et voilà que mon maître, qui vient de le consulter, s'en retourne plus fou que jamais ; il se laisse conduire par un aveugle, tout au contraire de ce qu'il devrait faire ; car c'est à ceux qui voient clair de guider les aveugles. Lui, il les suit, et il me force d'en faire autant ; et cela, sans même me dire le moindre mot. Non, mon maître ; je ne saurais me taire, si tu ne me dis pourquoi nous suivions cet homme : je te tourmenterai sans relâche, et tu ne me battras pas ; j'ai la couronne sur la tête.

Below the text windows, there are buttons for "CHRÉMYLE" and "CARION" in each window, and a "CHRÉMYLE" button at the bottom of the interface.



GenderedNews

Réalisation d'un dashboard pour visualiser les données

Karolin Boczoń (GRICAD)

Ange Richard (Pacte, LIG)

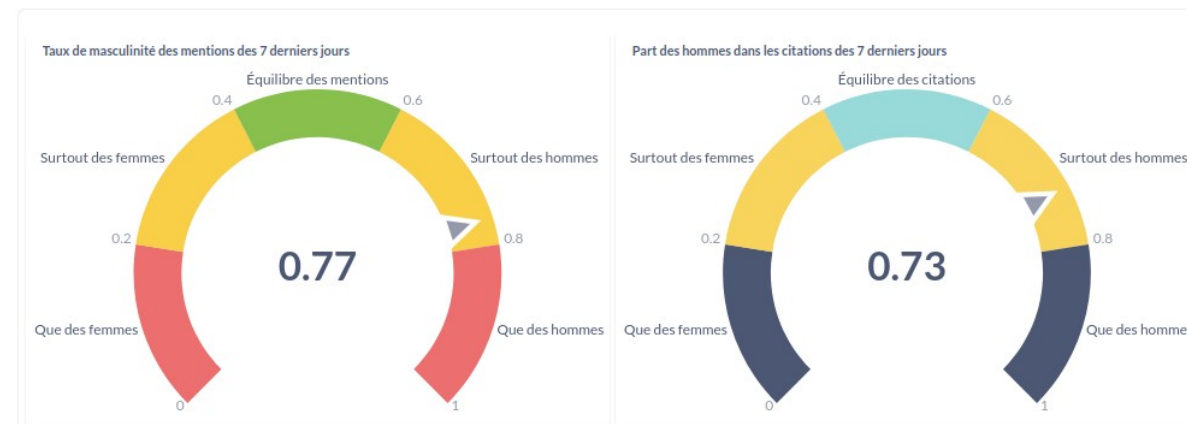
Le projet GenderedNews



GenderedNews mesure les inégalités de mention et de citation des hommes et des femmes dans les articles publiés par les principaux journaux français. Ces mesures permettent de rendre visible la part très importante des contenus des médias consacrée aux hommes, que ce soit en termes de mentions (le fait de parler d'hommes dans les articles) ou de citations (le fait de donner la parole à des hommes). Elles permettent aussi de visualiser les différences entre journaux et l'évolution de ces pratiques dans le temps.

Ce site web est actualisé tous les jours. Les résultats ci-dessous ne portent que sur les sept derniers jours. Pour avoir accès aux résultats détaillés depuis le début du projet, consultez les pages « Mesures » du site.

GenderedNews est un projet de recherche de l'[Université Grenoble Alpes](#). Il associe la [chaire Société algorithmique](#) et la chaire Intelligence Artificielle & Language de [MIAI](#).



Scraping et traitement quotidien de données du *web* (articles de journaux)

Deux indicateurs calculés sur le texte:

- Masculinité des mentions
- Masculinité des citations

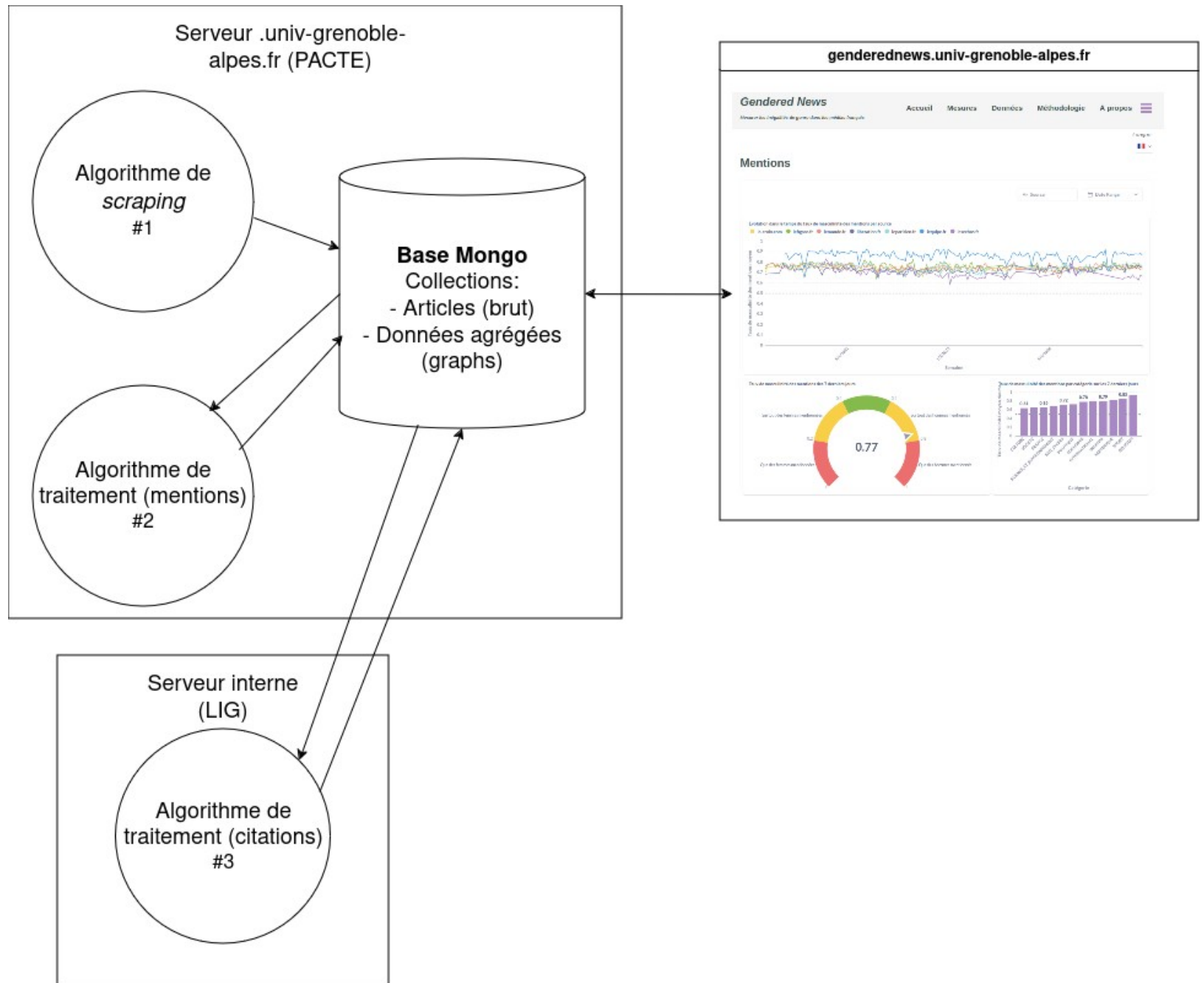
Données ouvertes:

- En 2021 : 7 journaux de la PQN
- Aujourd'hui : Une trentaine de sources

Besoin

Un *dashboard* existant à améliorer pour accommoder la nouvelle quantité de données

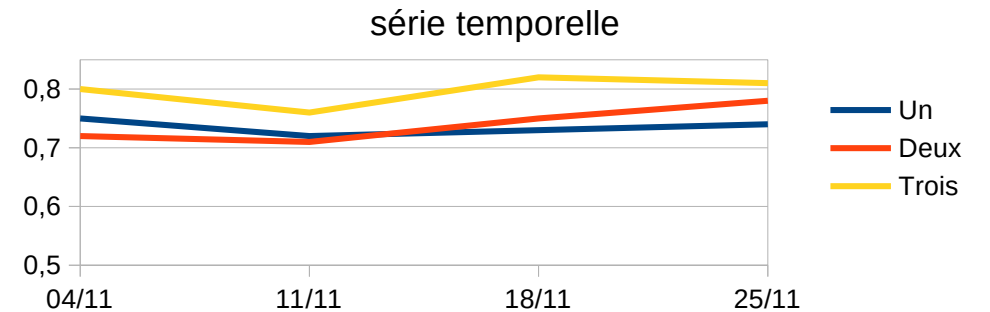
Tout en prenant en compte une architecture de traitement complexe



Visualisation des données

- Rend les données plus accessibles et compréhensibles.
- Permet d'identifier rapidement relations, tendances et modèles.
- Facilite l'accès aux informations pour un public non spécialiste.

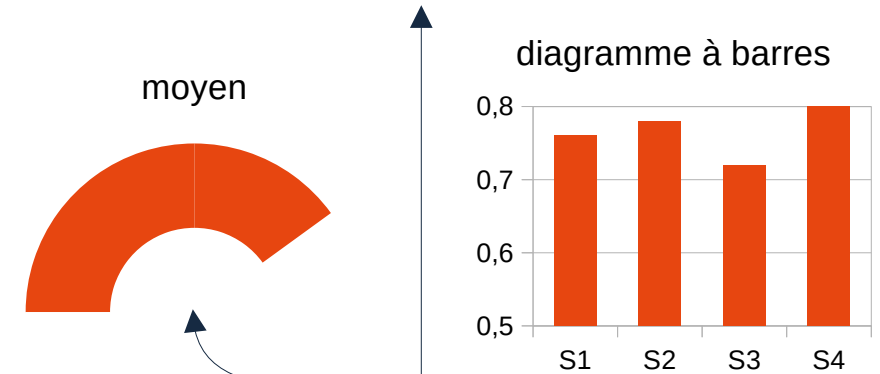
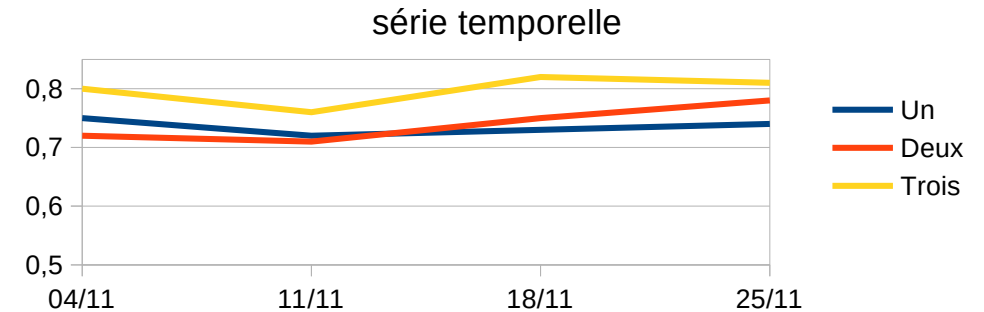
04/11/2024	0,75	0,72	0,8
11/11/2024	0,72	0,71	0,76
18/11/2024	0,73	0,75	0,82
25/11/2024	0,74	0,78	0,81



Dashboard

- Différents aspects des données peuvent être étudiés simultanément.
- Combiner plusieurs graphiques offre une vision plus complète.
- L'interactivité permet d'explorer les données selon ses besoins.

Sélectionner ▼



données

{id, date, periodique, sujet, taux}

Dash

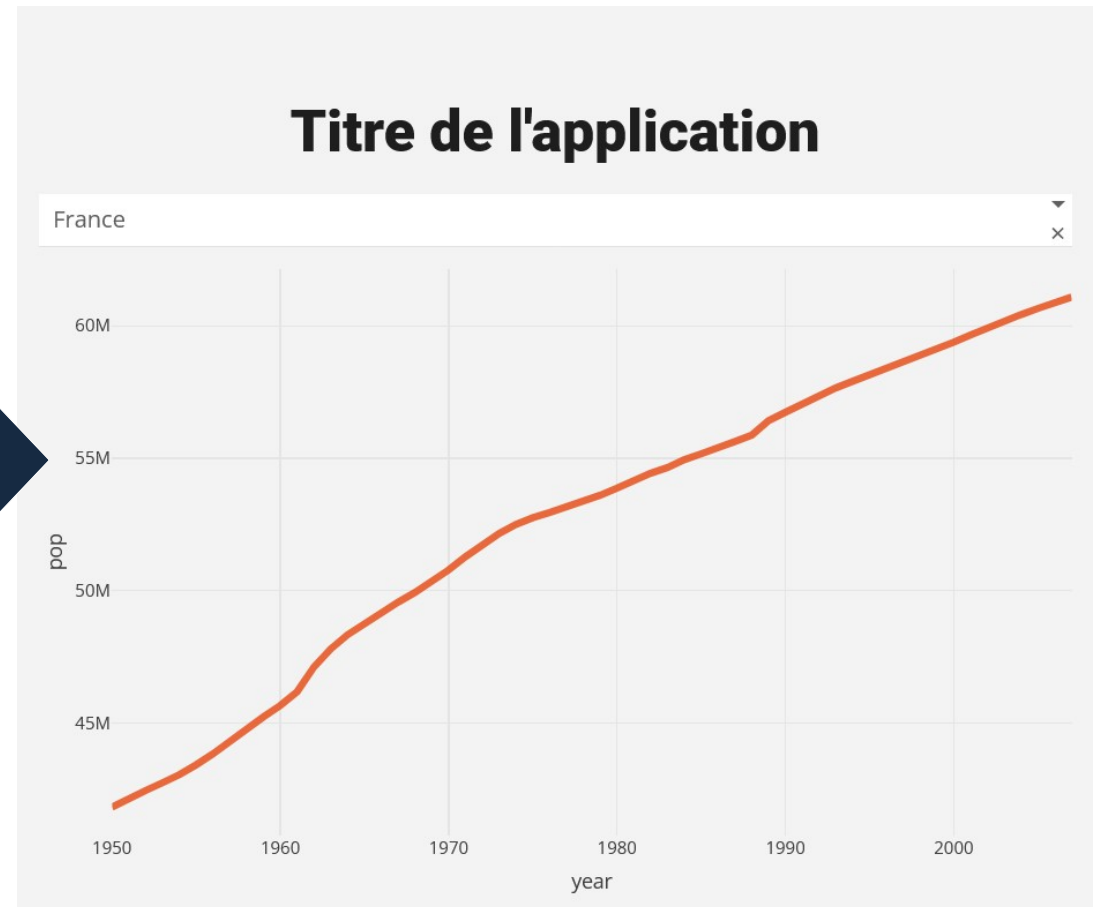
<https://dash.plotly.com/>

Dash est un framework pour créer rapidement des applications de données en *Python*.
La version de base est *open source*.

```
app = Dash()

app.layout = [
    html.H1(children='Titre de l'application',
             style={'textAlign': 'center'}),
    dcc.Dropdown(df.country.unique(), 'France',
                id='dropdown-selection'),
    dcc.Graph(id='graph-content')
]

@callback(
    Output('graph-content', 'figure'),
    Input('dropdown-selection', 'value')
)
def update_graph(value):
    dff = df[df.country==value]
    return px.line(dff, x='year', y='pop')
```



Dash

<https://dash.plotly.com/>

Réalisation de la mise en page avec "user interface" pour permettre à l'utilisateur d'explorer les données.

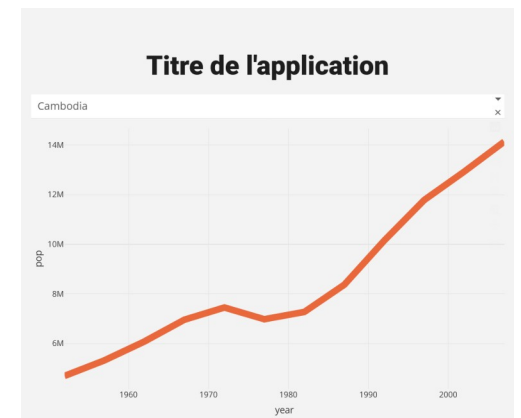
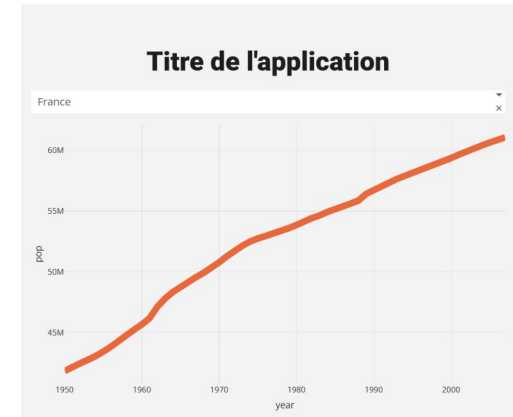
```
app = Dash()

app.layout = [
    html.H1(children='Titre de l'application',
             style={'textAlign': 'center'}),
    dcc.Dropdown(df.country.unique(), 'France',
                 id='dropdown-selection'),
    dcc.Graph(id='graph-content')
]

@callback(
    Output('graph-content', 'figure'),
    Input('dropdown-selection', 'value')
)
def update_graph(value):
    dff = df[df.country==value]
    return px.line(dff, x='year', y='pop')
```

disposition

interactivité



Cahier de charges

1. Site web

- l'en-tête, le pied de page, le menu de navigation, le bouton langue
- quatre sous-pages au total (y compris la page d'accueil)

2. Dashboard sur la page d'accueil

- interface de sélection du type de données, de la période de temps et des traces spécifiques

3. Graphiques

- série temporelle avec taux par le periodique dans le temps
- diagramme à barres avec taux moyen par sujet
- jauge avec le taux moyen total

Démonstration

gricad-gitlab.univ-grenoble-alpes.fr/boczonk/gates-genderednews

Gendered News

Mesurer les inégalités de genre dans les médias français

Accueil Données Méthodologie À propos

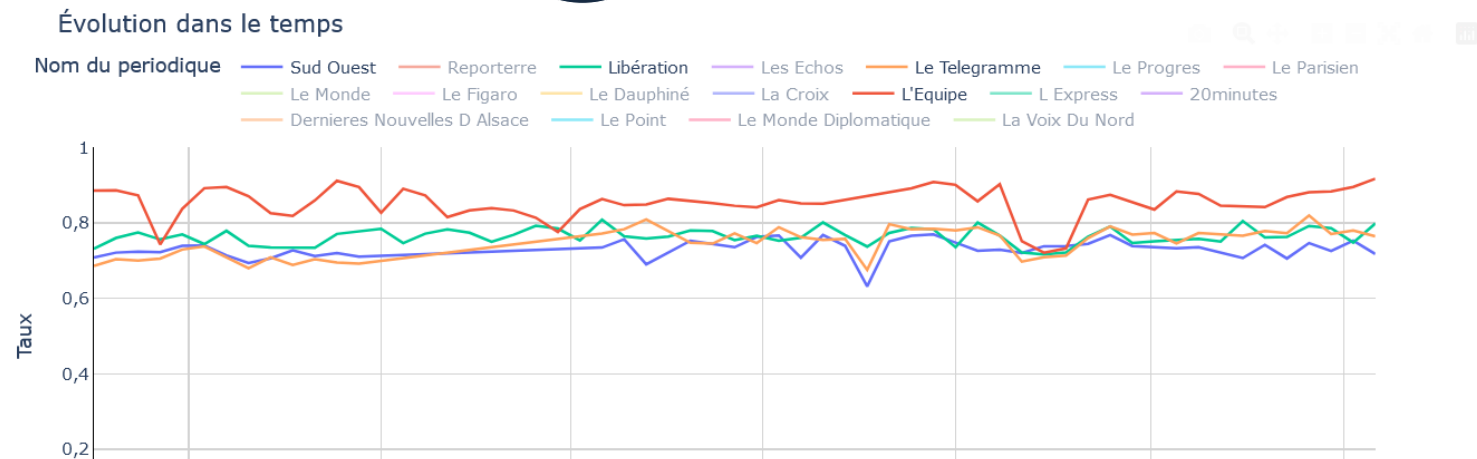
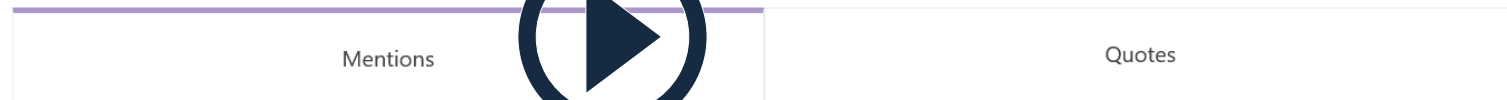
GenderedNews mesure les inégalités de mention et de citation des hommes et des femmes dans les articles publiés par les principaux journaux français. Ces mesures permettent de rendre visible la part très importante des contenus des médias consacrée aux hommes, que ce soit en termes de mentions (le fait de parler d'hommes dans les articles) ou de citations (le fait de donner la parole à des hommes). Elles permettent aussi de visualiser les différences entre journaux et l'évolution de ces pratiques dans le temps.

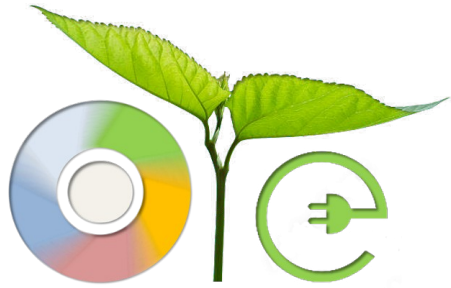
GenderedNews est un projet de recherche de l'[Université Grenoble Alpes](#). Il associe la [chaire Société algorithmique](#) et la chaire Intelligence Artificielle & Language de [MIAI](#).

Modifiez les dates et sélectionnez les sources à l'aide des menus ci-dessous. Sélectionnez une zone avec la souris pour zoomer. Double-cliquer pour revenir.

02/10/2023 →
11/11/2024

- × Sud Ouest
- × Libération
- × Le Telegramme
- × L'Equipe





Observatoire de la Transition Énergétique

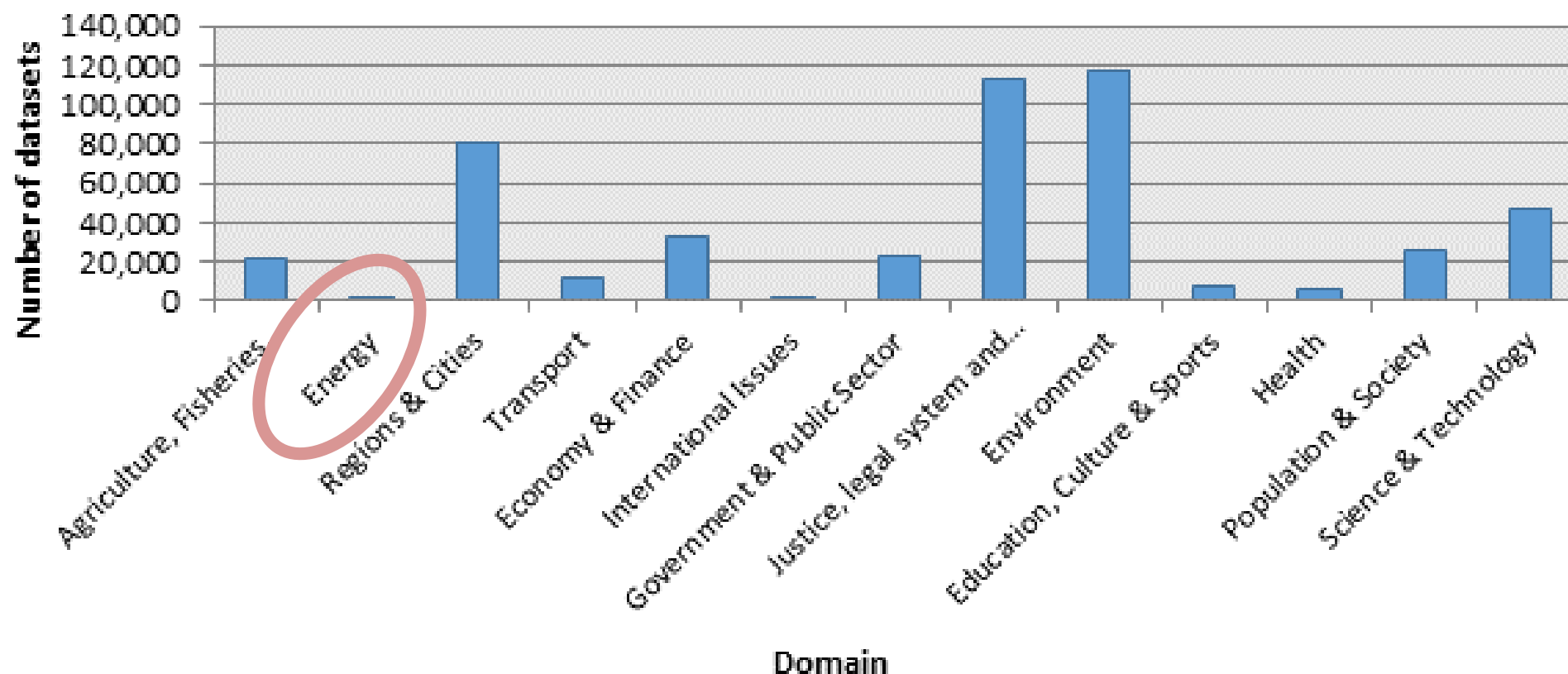
Christophe Boisseau (OTE)
Lilas Lacoste (GRICAD)

I'OTE pourquoi ?

- Urgence climatique et environnementale qui s'accélère
- 75 % des impacts CO2 sont dus à des activités énergétiques
- Consensus scientifique :
 - Il faut décarboner
 - L'électrification massive des usages est nécessaire
 - Il faut de l'efficacité
 - Mais aussi de la sobriété et de la flexibilité : mobilisation de leviers sociaux-techniques
- Écart entre ce consensus et la réalité des perceptions, des opinions, des avis, des débats et controverses
- Comment faire la jonction ?

I'OTE pourquoi ?

Datasets per category

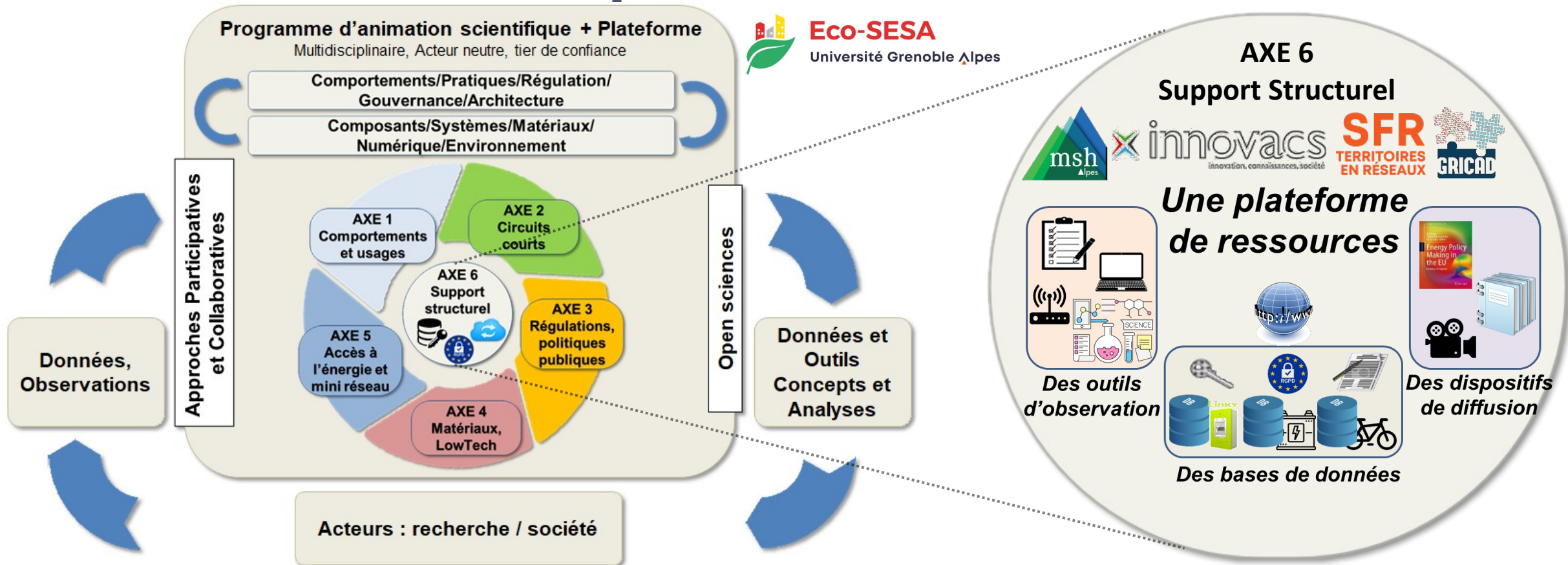


I'OTE c'est quoi ?

Une synergie entre un programme scientifique et une plate-forme de ressources mutualisées

La science ouverte et participative au service de la transition

L'OTE c'est quoi ?



104
Chercheuses et chercheurs



19
Laboratoires


UGA
Université Grenoble Alpes

11
Domaines scientifiques



l'OTE comment ? Pourquoi ?

Le recueil et l'accès aux données « propriétaires » d'acteurs tiers


Les individus : RGPD
- Nécessité d'un consentement éclairé
- Données disponibles pour un temps limité

Les organisations, personnes morales (communautés, coopératives, territoires, entreprises, ...)
- Nécessité d'un accord de mise à disposition

Avoir une structure de type « tiers de confiance »

Apte et légitime à recueillir les données des **tiers**

Pour produire des données interdisciplinaires de la recherche

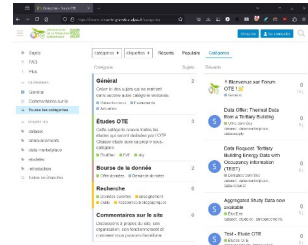
Anticiper – accompagner – pérenniser

- publications en open-access
- données en open-data
- codes en open-source
- dispositifs en open-hardware

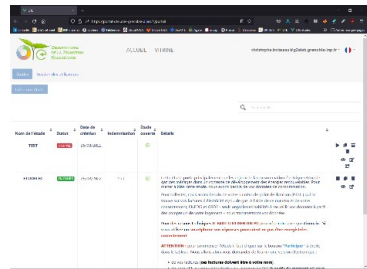
L'OTE comment ?



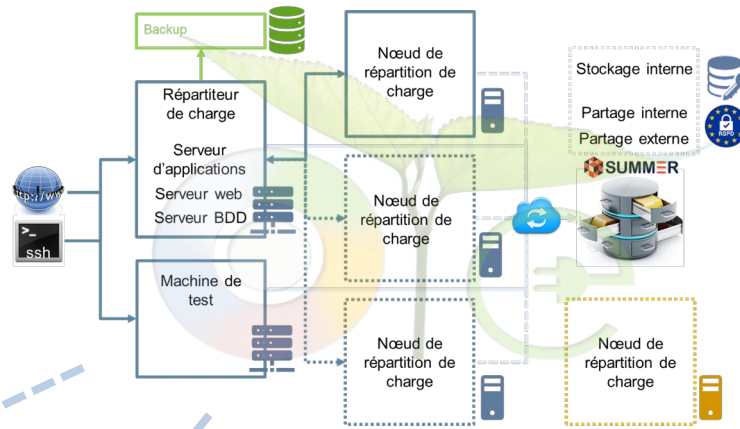
<https://ote.univ-grenoble-alpes.fr>



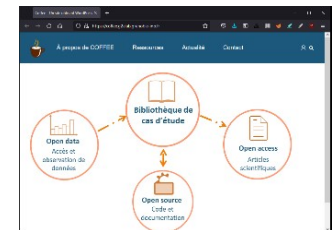
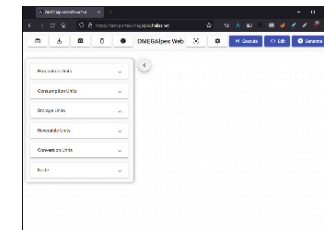
<https://forum-ote.univ-grenoble-alpes.fr/>



<https://portal-ote.univ-grenoble-alpes.fr/studies>



Serveur de
Fichiers
(Protocoles
SMB et FTP)



<https://omegalpes-ote.univ-grenoble-alpes.fr>

<https://coffee-ote.univ-grenoble-alpes.fr>

Sur plateforme de test...



<https://jupyter-n1ote.univ-grenoble-alpes.fr>



Déploiements à l'étude...



<https://influxdb-ote.univ-grenoble-alpes.fr>

<https://grafana-ote.univ-grenoble-alpes.fr>

L'OTE comment ?

- Adhésion à l'OTE

- <https://ote.univ-grenoble-alpes.fr/adherez/>
- Questionnaire générique
→ Panel OTE

- Accès aux études

- <https://portal-ote.univ-grenoble-alpes.fr/login>



Adhérer - Observatoire de la T. X

https://ote.univ-grenoble-alpes.fr/adherez/

Observatoire de la Transition Énergétique
La science ouverte et participative au service de la transition

Accueil Actualité **Participez** Qui sommes nous ? Projets OTE FlexTASE Résultats L'OTE en temps réel

REGISTER NOW

Si vous souhaitez participer et aider les chercheurs qui travaillent sur les problématiques liées à la transition énergétique vous pouvez adhérer à l'OTE.
Concrètement cette adhésion se résume à répondre à un questionnaire sommaire. Une fois inscrit vous pourrez :

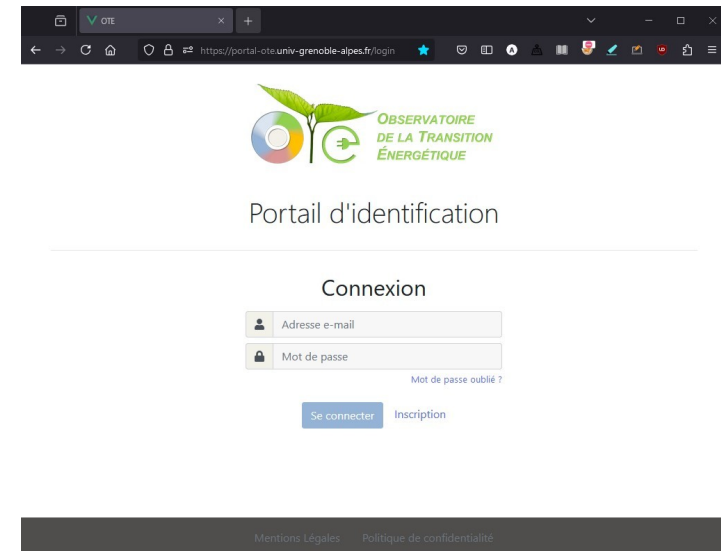
- soit être contactés par des chercheurs pour vous proposer des études liées à la transition,
- soit vous porter volontaire pour les études en cours.

[Formulaire d'inscription](#)

Si vous êtes déjà inscrit à l'OTE et que vous voulez participer à une nouvelle étude [connectez-vous](#).

Rechercher

Assemblée générale FlexTASE
par OTEAdmin
28 novembre 2024



OTE

https://portal-ote.univ-grenoble-alpes.fr/login

Observatoire de la Transition Énergétique

Portail d'identification

Connexion

Adresse e-mail

Mot de passe

Mot de passe oublié ?

Se connecter Inscription

Mentions Légales Politique de confidentialité

Besoins exprimés et réalisation (1) : mise en forme de données

Mise en forme de données d'enquête

Format de données des enquêtes inexploitable pour les analyses

Avez-vous participé à notre dernière étude [...] ? Oui / Non

Si Oui :

Situation 1 : est-ce que vous préférez l'option A ou l'option B ? Option A /Option B

Situation 2 : est-ce que vous préférez l'option A ou l'option B ? Option A /Option B

....

Si Non :

Exemple de format des données brutes

ID	Avez-vous participé à notre dernière étude [...] ?	Oui	Situation 1 : est-ce que vous préférez l'option A ou l'option B ?	Option A	Option B	Situation 2 : est-ce que vous préférez l'option A ou l'option B ?	Option A	Option B	...	Non	...
repondant_1		X		X				X			
repondant_2										X	

→ **Besoin de pouvoir identifier clairement chaque question et ses réponses dans les données !**

→ **Mise en forme des données brutes nettoyage, renommage, rassemblement de données, recodage...**

ID	Q1	Q2	Q3	...
repondant_1	1	1	2	
repondant_2	2	99	99	



Technologies :
python (pandas)

→ script : **reproductibilité des traitements** si ajout de nouvelles données

→ **Génération automatique d'un dictionnaire des variables**

Question	Enoncé	Type de question	Modalité de réponses	Barème
Q1	Avez-vous participé à notre dernière étude de l'Observatoire de la Transition Énergétique (Etudelec) au mois d'aout/septembre ?	fermée à choix unique	"Oui", "Non", non-réponse	[1,2,99]
Q2	Situation 1 : est-ce que vous préférez l'option A ou l'option B ?	fermée à choix unique	"Option A", "Option B", non-réponse	[1,2,99]
Q3	Situation 2 : est-ce que vous préférez l'option A ou l'option B ?	fermée à choix unique	"Option A", "Option B", non-réponse	[1,2,99]

Besoins exprimés et réalisation (2) : datavisualisation

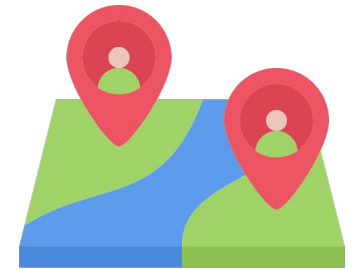
Visualisation :

- **Répartition géographique** des panélistes à différentes échelles (communes, départements..)
- **Consommation** électrique moyenne des panélistes **par zone géographique** (régions, départements..)

Observatoire de la transition énergétique

Répartition géographique des panélistes à différentes échelles (communes, départements..)

Données panel :
**code postal et
noms de
communes**



Création de cartes

Observatoire de la transition énergétique

Répartition géographique des panélistes à différentes échelles (communes, EPCI..)



geo.api.gouv.fr

Données panel :
code postal et
noms de
communes



Recherche
**données de
découpage
administratif** :
données
ouvertes, à jour,
documentées



Création de cartes

Observatoire de la transition énergétique

Recherche **données de découpage administratif** : données ouvertes, à jour, documentées

→ pour chaque commune, avoir **nom et code** des département, région de rattachement..

→ **code unique** pour chaque entité : code INSEE commune ...

→ **contour géographique** de chaque entité

→ éléments de jointure avec les données du panel : **code postal + nom de commune**

Observatoire de la transition énergétique

Recherche **données de découpage administratif** : données ouvertes, à jour, documentées

Code Officiel Géographique (COG) INSEE

Insee Institut national de la statistique et des études économiques

Menu Blog Presse Aide English

Rechercher

STATISTIQUES ET ÉTUDES DÉFINITIONS, MÉTHODES ET QUALITÉ SERVICES L'INSEE ET LA STATISTIQUE PUBLIQUE

Accueil > Téléchargement de fichiers > Géographie administrative > Code officiel géographique (COG) > Code officiel géographique au 1^{er} janvier 2024

Code officiel géographique au 1^{er} janvier 2024

Téléchargement des fichiers

GÉOGRAPHIE
Date de publication : 20/02/2024
> Imprimer

ENSEMBLE DES FICHIERS 2024 (CSV)
(zip, 1 Mo)

ENSEMBLE DES FICHIERS 2024 (DBASE)
(zip, 2 Mo)

- **code unique** ✓
- **contour géographique** X
- **code postal** + **nom de commune** X

Communes

Liste des communes, arrondissements municipaux, communes déléguées et communes associées au 1^{er} janvier 2024, avec le code des niveaux supérieurs (canton ou pseudo-canton, département, région)

Liste des communes

(dbase, 6 Mo) (csv, 3 Mo)

- Dessin du fichier « Communes existantes » et liste des variables

Nom de la variable	Longueur de la variable	Désignation et modalités de la variable
TYPECOM	4	Type de commune
COM	5	Code commune
REG	2	Code région
DEP	3	Code département
CTCD	4	Code de la collectivité territoriale ayant les compétences départementales
ARR	4	Code arrondissement

Observatoire de la transition énergétique

Recherche **données de découpage administratif** : données ouvertes, à jour, documentées

DATANOVA La Poste – base officielle des codes postaux



The screenshot shows a web browser window with the URL datanova.laposte.fr/datasets/laposte-hexasmal/full. The page features the LA POSTE GROUPE logo and a navigation breadcrumb: Accueil > Données > Base officielle des codes postaux > Vue tabulaire. A search bar contains the text "Rechercher". Below the search bar, a table displays 39,192 rows of data. The table has six columns: #Code commune INSEE, Nom de la commune, Code postal, Libellé d'acheminement, Lieu dit - nom de la commune associée, and geometry. The first five rows of the table are visible.

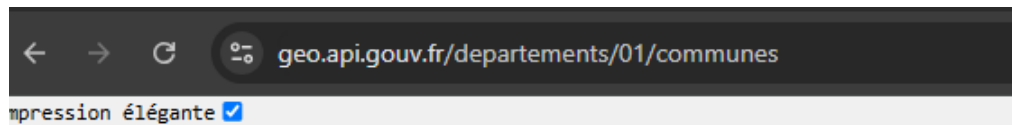
#Code commune INSEE	Nom de la commune	Code postal	Libellé d'acheminement	Lieu dit - nom de la commune associée	geometry
01001	L ABERGEMENT CLEMENCIAT	01400	L ABERGEMENT CLEMENCIAT		{"type":"Polygon","coordinates":[[[4.9584116,46.15.
01002	L ABERGEMENT DE VAREY	01640	L ABERGEMENT DE VAREY		{"type":"Polygon","coordinates":[[[5.4302024,45.98.
01004	AMBERIEU EN BUGEY	01500	AMBERIEU EN BUGEY		{"type":"Polygon","coordinates":[[[5.4088198,45.94.
01005	AMBERIEUX EN DOMBES	01330	AMBERIEUX EN DOMBES		{"type":"Polygon","coordinates":[[[4.9429765,45.97.
01006	AMBLEON	01300	AMBLEON		{"type":"Polygon","coordinates":[[[5.5708254,45.75.

- **code unique** ✓
- **contour géographique** ✓
- **code postal** + **nom de commune (format)** X

Observatoire de la transition énergétique

Recherche **données de découpage administratif** : données ouvertes, à jour, documentées

Geo.api.gouv



```
{
  "nom": "L'Abergement-Clémenciat",
  "code": "01001",
  "codeDepartement": "01",
  "siren": "210100012",
  "codeEpci": "200069193",
  "codeRegion": "84",
  "codesPostaux": [
    "01400"
  ],
  "population": 832
},
{
  "nom": "L'Abergement-de-Varey",
  "code": "01002",
  "codeDepartement": "01",
```

- **code unique** ✓
- **contour géographique** ✓
- **code postal + nom de commune** ✓

Documentation technique

GET /communes Rechercher des communes

Nom	Description	Type
codePostal	Code postal associé	string
lon	Longitude (en degrés)	number
lat	Latitude (en degrés)	number
nom	Nom de la commune	string
boost	Mode de boost de la recherche par nom	string
code	Code (INSEE) de la commune	string
siren	Code SIREN de la commune	string
codeEpci	Code de l'EPCI associé	string
codeDepartement	Code du département associé	string
codeRegion	Code de la région associée	string
zone	Zone, métropole (metro), DROM (drom) et COM (com)	array [string]
type	Type, peut être commune-actuelle et/ou arrondissement-municipal	array [string]
fields	Liste des champs souhaités dans la réponse	array [string]
format	Format de réponse attendu	string

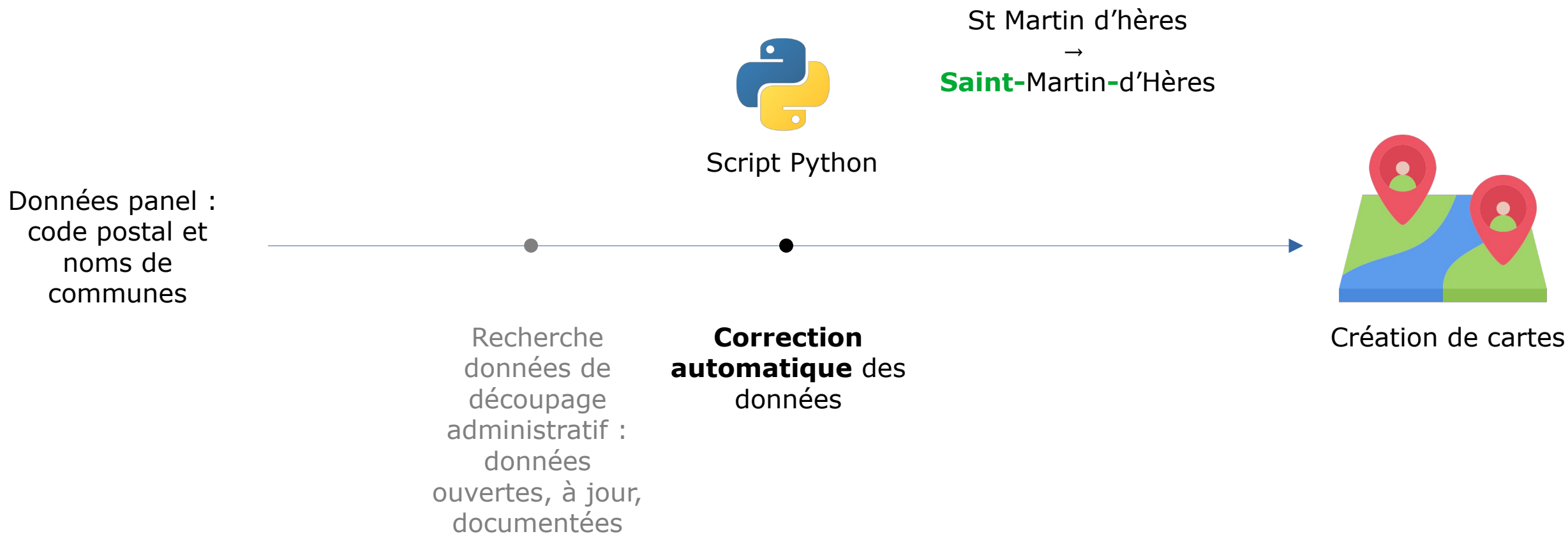
json | **geojson**

geometry Géométrie à utiliser pour la sortie géographique (format GeoJSON) string

centre | **contour** | bbox | mairie

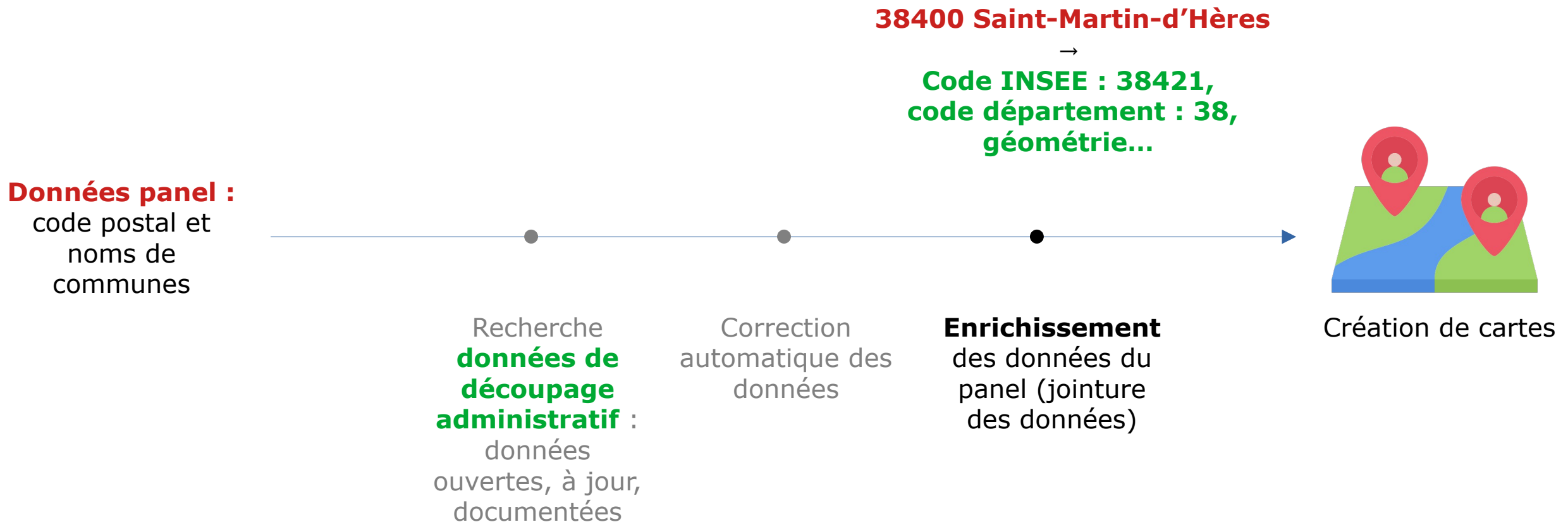
Observatoire de la transition énergétique

Répartition géographique des panélistes à différentes échelles (communes, EPCI..)



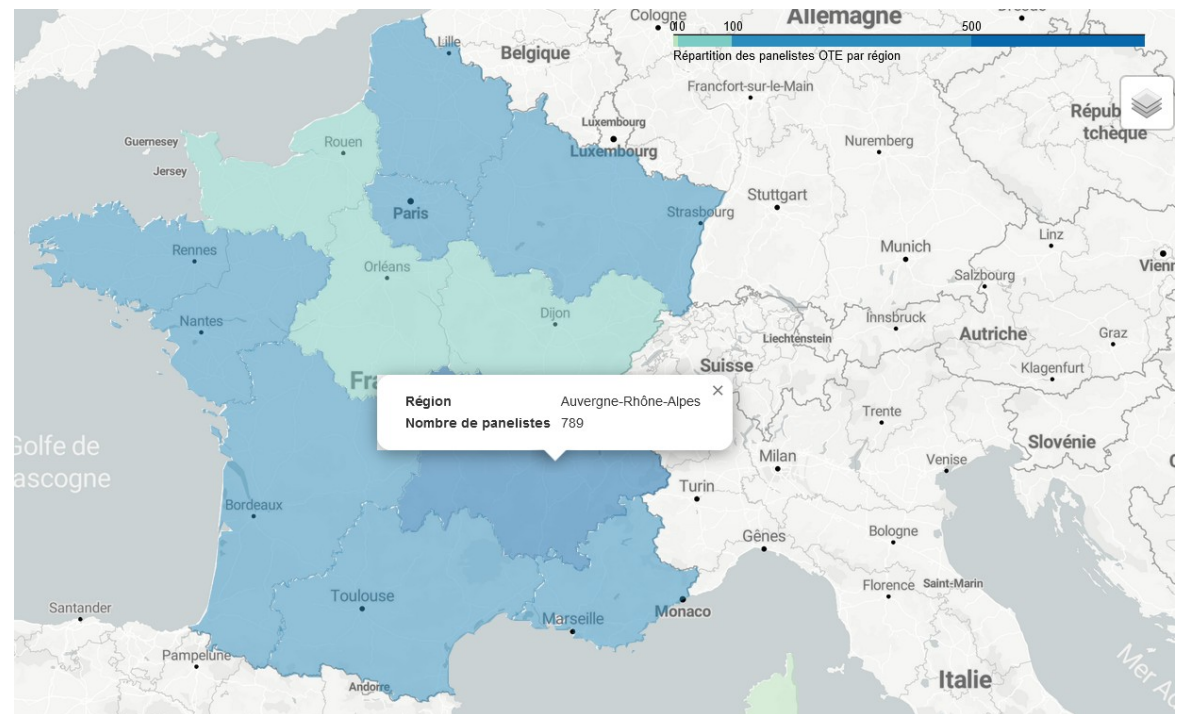
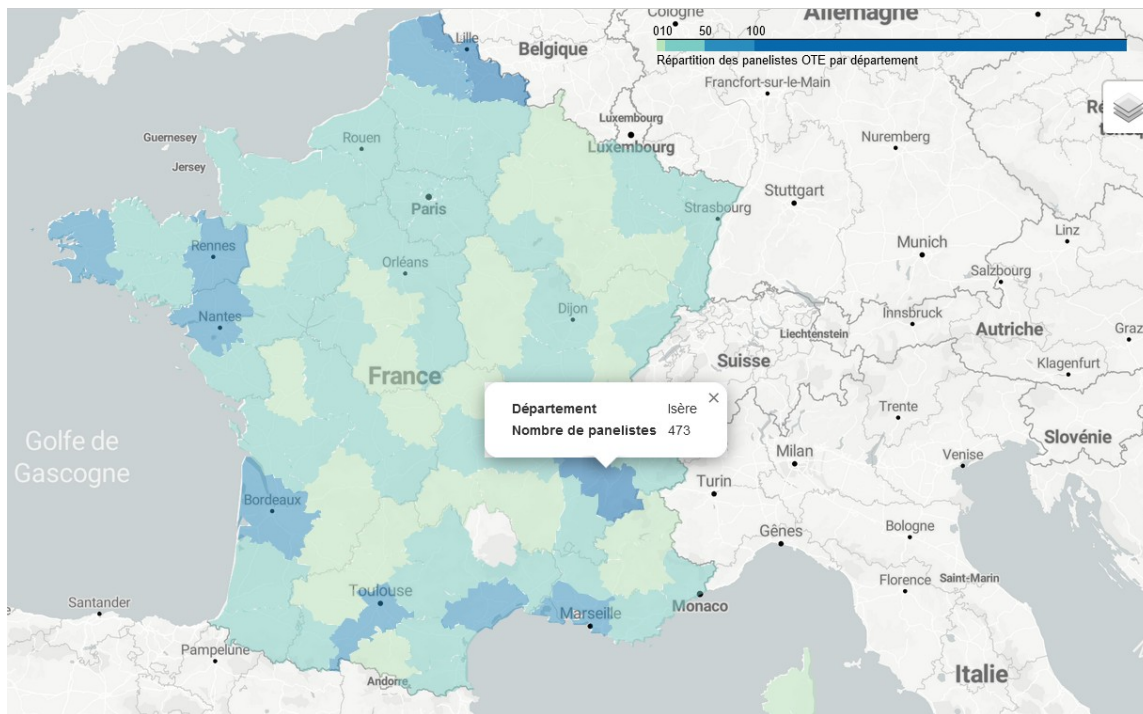
Observatoire de la transition énergétique

Répartition géographique des panélistes à différentes échelles (communes, EPCI..)



Observatoire de la transition énergétique

Répartition géographique des panélistes à différentes échelles (communes, EPCI..)



Technologies :
Python (folium, pandas, geopandas)
Leaflet




Observatoire de la transition énergétique

Consommation électrique moyenne des panélistes **par zone géographique** (régions, départements..)

Étape préliminaire de **traitement des données** de consommation électrique **manquantes**

→ **Complétion**

dates	conso_sujet_312
12/01/2023 11:00:00	216
12/01/2023 11:30:00	180
12/01/2023 12:00:00	
12/01/2023 11:30:00	195

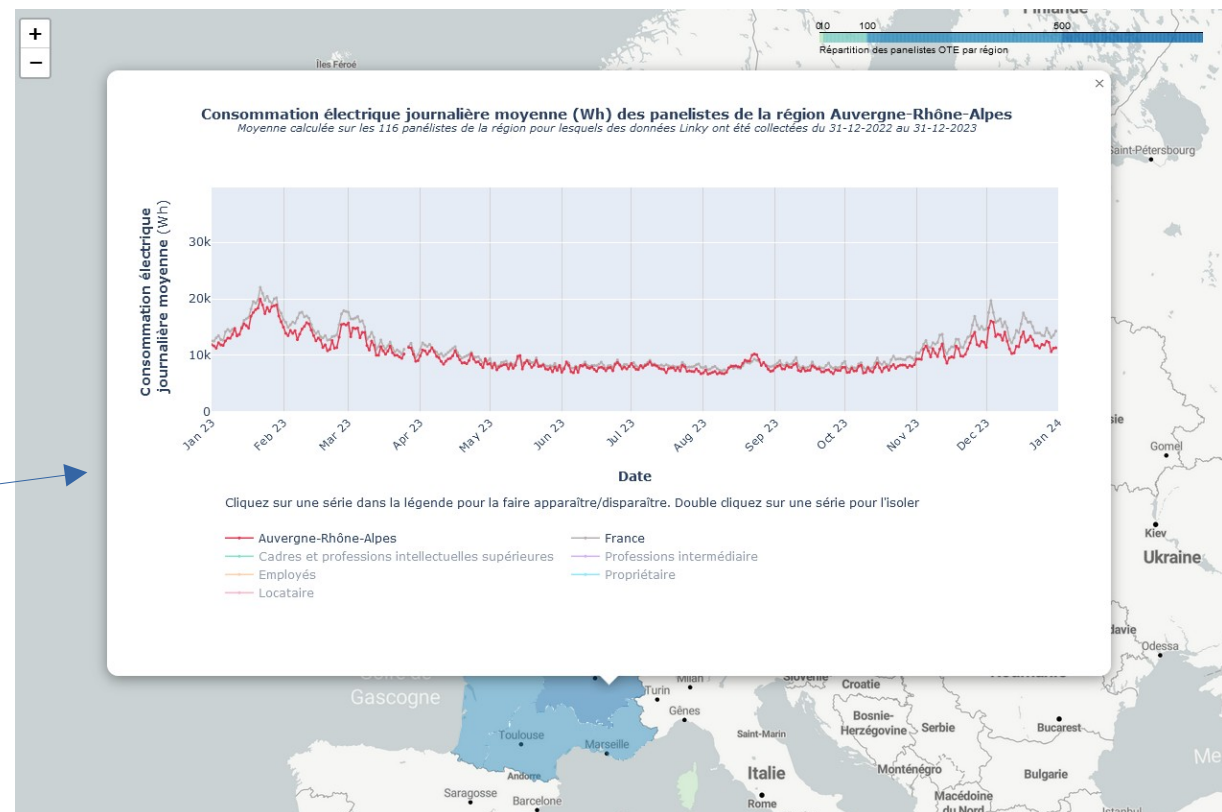
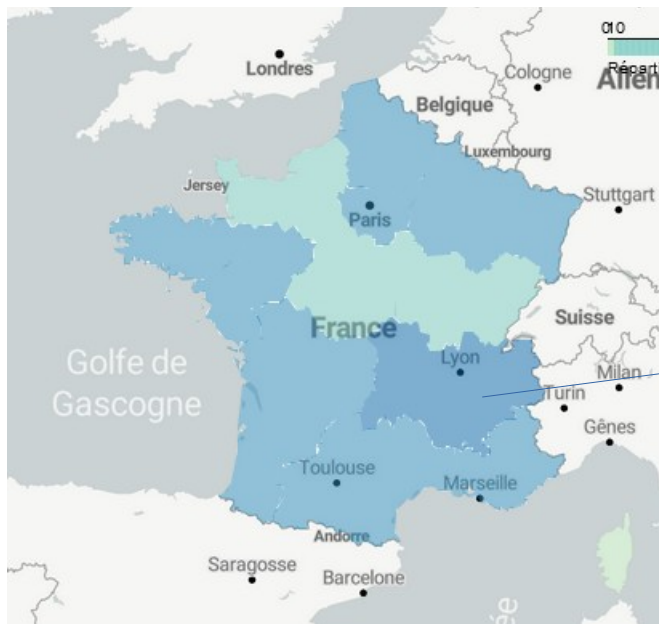


dates	conso_sujet_312
12/01/2023 11:00:00	216
12/01/2023 11:30:00	180
12/01/2023 12:00:00	180
12/01/2023 11:30:00	195

→ **Exclusion** des journées si trop de données manquantes

Observatoire de la transition énergétique

Consommation électrique moyenne des panélistes **par zone géographique** (régions, départements..)



<https://ote.univ-grenoble-alpes.fr/courbes-de-charge-panel-ote/>

I'OTE valorisation

- Open data : Recherche Data Gouv
 - Collection OTE : <https://entrepot.recherche.data.gouv.fr/dataverse/ote>
 - Publication des jeux de données :
 - ExpeSIGNO : <https://doi.org/10.57745/YTSBGL>
 - EtudElec : <https://doi.org/10.57745/WIWMMK>

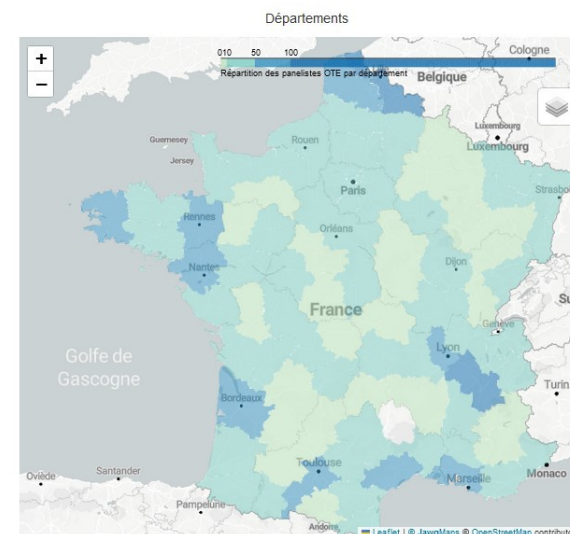
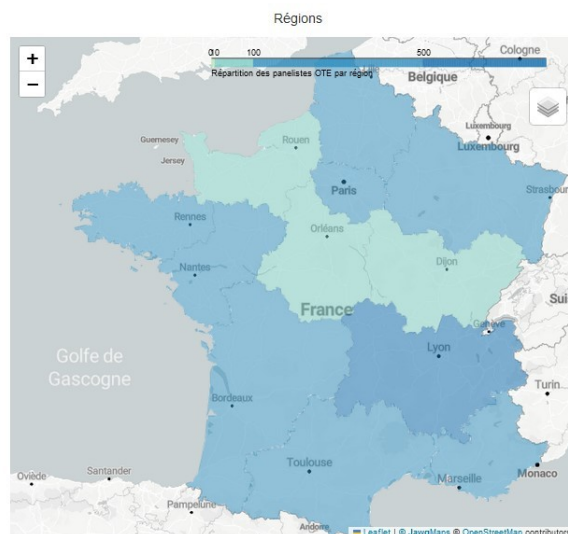
l'OTE valorisation

- Géographie du panel
- 4 niveaux géographiques : <https://ote.univ-grenoble-alpes.fr/geographie-du-panel-ote/>

Représentation géographique du panel de l'OTE

Nous vous proposons dans cette page une représentation géographique du panel de l'OTE. Les cartes ci-dessous sont interactives (en passant la souris ou en cliquant sur une zone elles indiquent le nombre de « panélistes » compris dans la zone). Elles peuvent mettre quelques secondes à s'afficher et sont mises à jour automatiquement toutes les semaines.

Nombre de « panélistes » par :

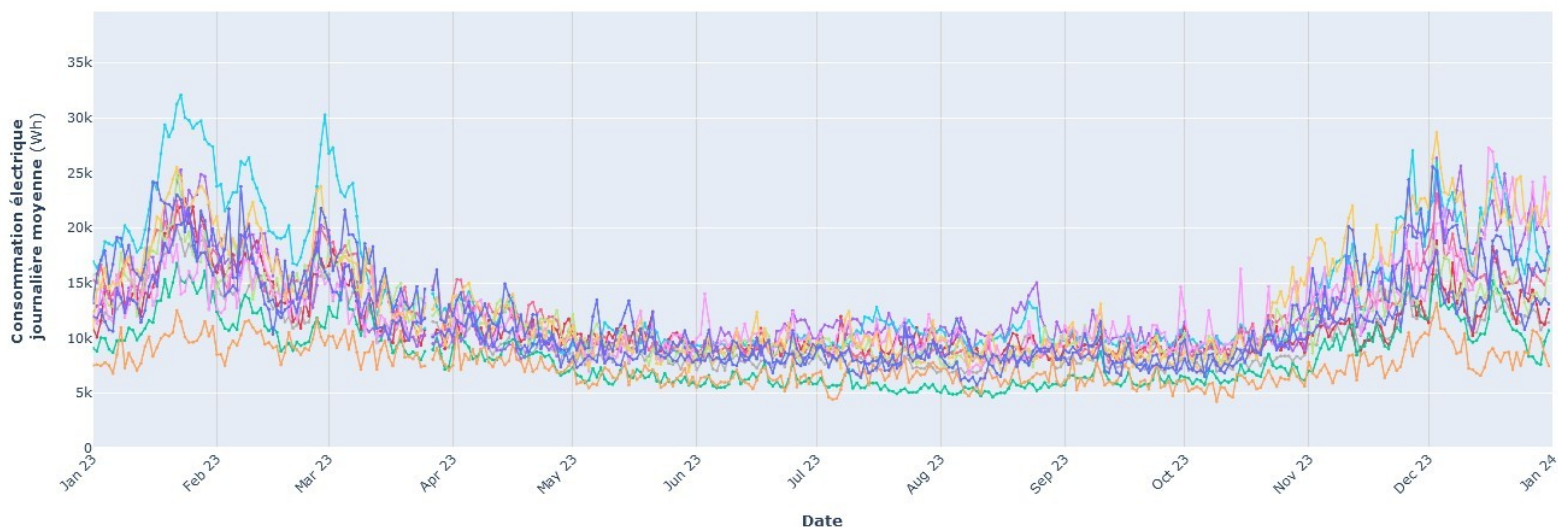


l'OTE valorisation

- Courbes de charges du panel
- Avec différents filtres : <https://ote.univ-grenoble-alpes.fr/courbes-de-charge-panel-ote/>

Consommation journalière moyenne par régions sur l'année 2023 :

Consommation électrique journalière moyenne (Wh) des panelistes des différentes régions
Moyenne calculée sur les 385 panelistes de la région pour lesquels des données Linky ont été collectées du 31-12-2022 au 31-12-2023



Cliquez sur une série dans la légende pour la faire apparaître/disparaître. Double cliquez sur une série pour l'isoler

— Auvergne-Rhône-Alpes	— Nouvelle-Aquitaine	— Île-de-France	— Provence-Alpes-Côte d'Azur	— Grand Est
— Occitanie	— Hauts-de-France	— Bretagne	— Centre-Val de Loire	— Bourgogne-Franche-Comté
— Pays de la Loire	— Normandie			

Projet IDEX Justice Algorithmique des élections (JADE) 2022-2025

Romain Rambaud, Professeur de Droit Public, CRJ
Louis Maritaud, Ingénieur de Recherche GATES

Description du projet

- **Constat** : montée en puissance de la « justice prédictive », de la « jurimétrie » mais peu de travaux en droit public, volonté de développer un projet en IA
- **Hypothèse / contentieux des élections politiques** : contentieux ancien, standardisé, normé, chiffré, donc plus réceptif aux outils de la justice algorithmique.... **Problématique** : quelle rationalité = notion de sincérité du scrutin, écart de voix ? (v. R. Rambaud, AJDA)
- **Méthodologie** : hypothèse qu'une « méthodologie » de la justice algorithmique se développe MAIS qu'elle est actuellement surtout développée par des start-ups / Legal tech donc échappent à la recherche publique et au service public (thèse A. Hafsaoui).

Élaboration technique des bases de données JADE

- Les bases de données Décisions : la problématique de la (non)préparation des données / métadonnées par les juridictions (XML – Zip Datagouv)

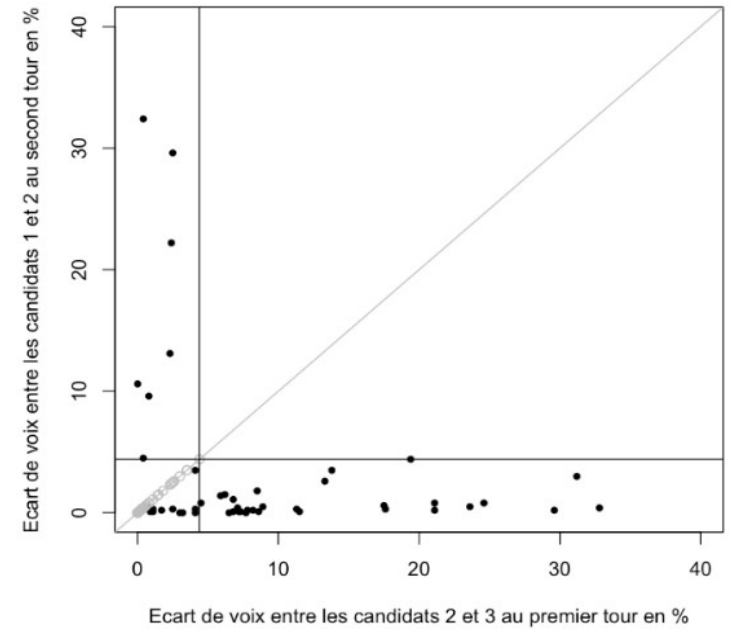
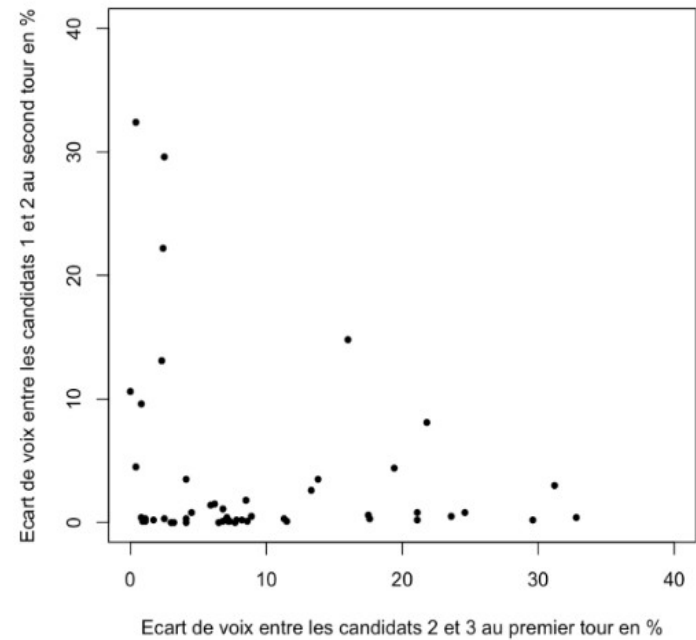
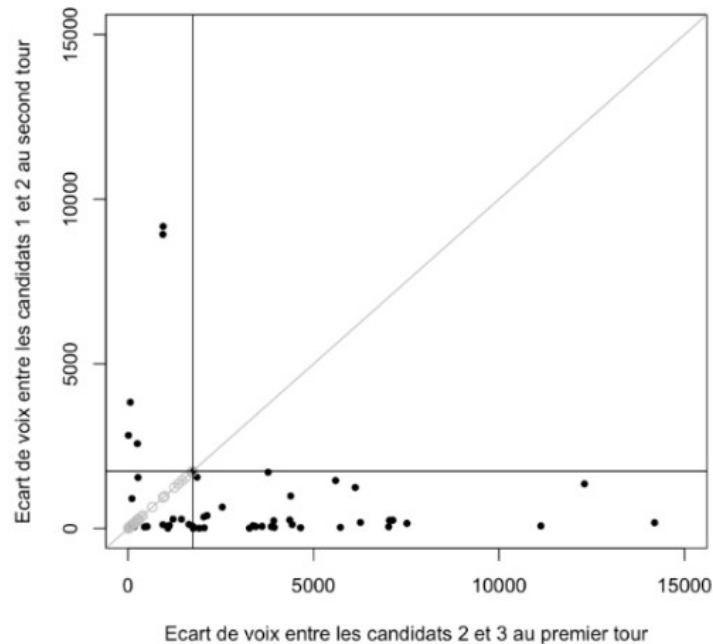
- **Conseil constitutionnel** : Les balises sémantiques suivantes permettent de trier informatiquement les décisions selon : <ECLI> (numéro unique de la décision), <URL> (lien internet), <NATURE> (AN, SEN, etc.), <TITRE> (type d'élection et circonscription concernée), <DATE_DEC> (date) et <SOLUTION> (sens de la décision). Il existe en plus sur certains décisions les « Abstracts » en JSON

- **Conseil d'Etat** :

Code PJCA	Texte Code PCJA	Total décisions	Balises solution remplies			Balises solution non remplies	
			Nombre	%	Nb Métadonnées	Nombre	%
28-005	Dispositions générales	240	98	40,8	25	142	59,2
28-01	Présidentielle	13	7	53,8	5	6	46,2
28-02	Législatives	42	23	54,8	7	19	45,2
28-023	Européennes	37	10	27,0	3	27	73,0
28-025 sauf 28-025-04	Régionales	176	29	16,5	10	147	83,5
28-03	Départementales	465	122	26,2	23	343	73,8
28-04	Municipales	1868	470	25,2	61	1398	74,8
28-08	Procédure spéciale	820	359	43,8	104	461	56,2

Résultats de JADE : des statistiques descriptives...

1) La prise en compte par le juge d'un écart de voix faible entre les candidats 2 et 3 du 1^{er} tour ou entre les candidats 1 et 2 du 2nd tour



2) Analyse du faible écart de voix pris en compte par le juge de l'élection

Indicateur statistique	Valeur Absolue	% des suffrages exprimés
Minimum	5	0
Premier quartile (25%)	56	0,1
Médiane	119	0,4
Troisième quartile (75%)	281	1,4
Extrémité supérieure de la moustache	394	3,0

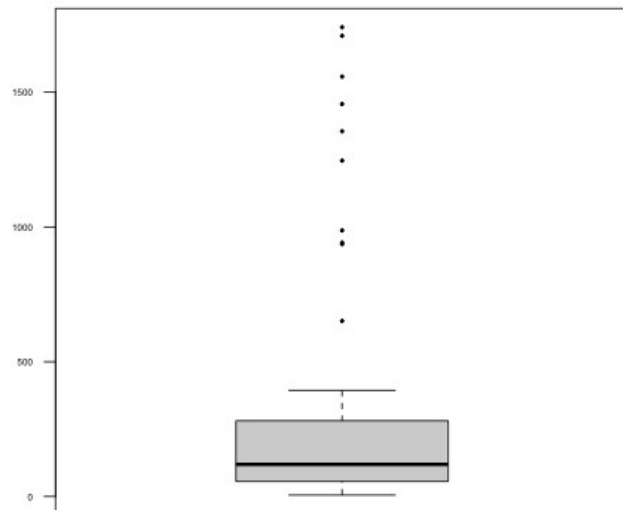


Figure 4 : Plus petits écarts de voix entre les candidats 2 et 3 du tour 1 et les candidats 1 et 2 du tour 2 en valeur absolue

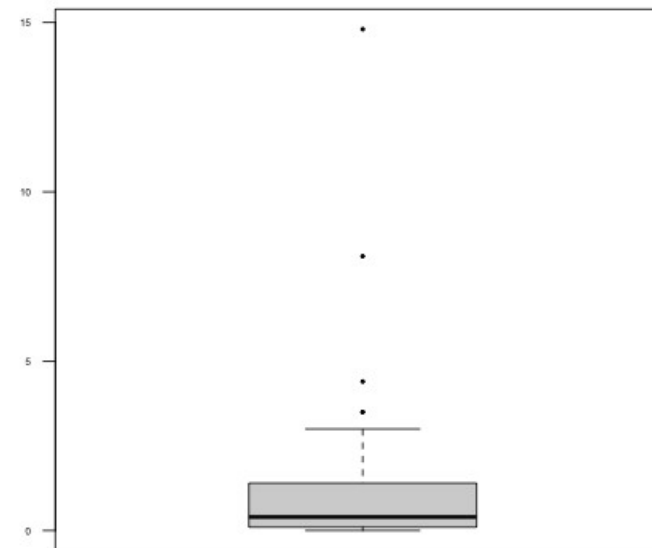


Figure 5 : Plus petits écarts de voix entre les candidats 2 et 3 du tour 1 et les candidats 1 et 2 du tour 2 en %

3) Hypothèse sur l'évolution dans le temps de l'appréciation de l'écart de voix

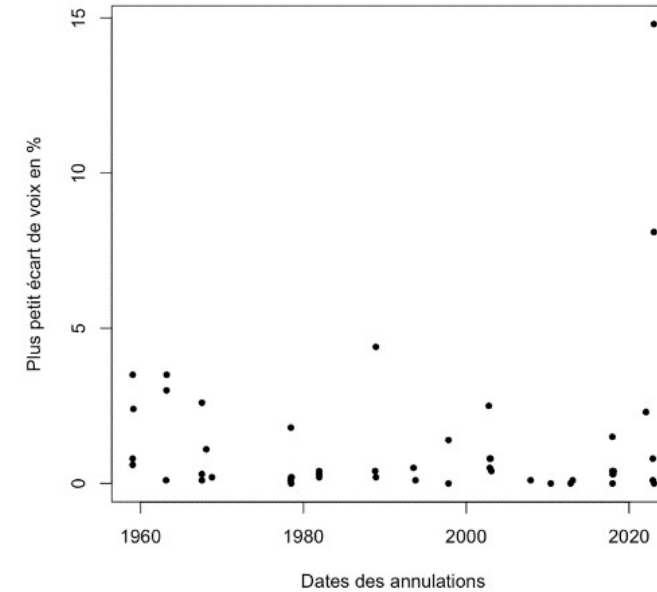
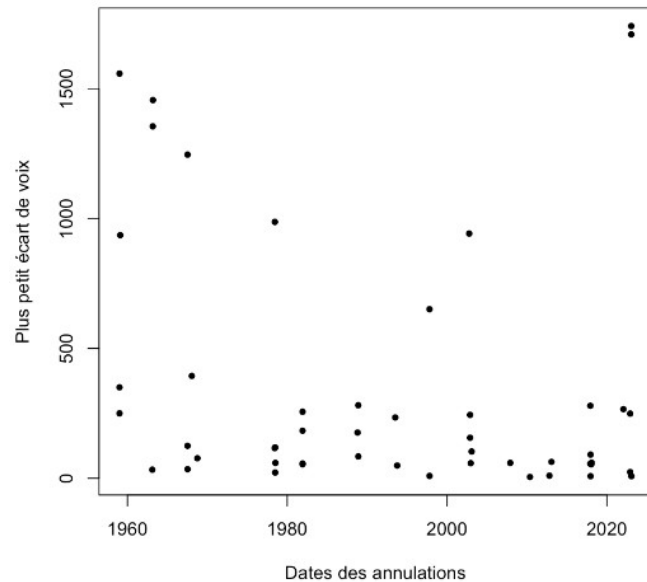


Figure 7 : Plus petit écart de voix entre les candidats 2 et 3 du tour 1 et les candidats 1 et 2 du tour 2 en pourcentage au cours du temps

Figure 6 : Plus petit écart de voix entre les candidats 2 et 3 du tour 1 et les candidats 1 et 2 du tour 2 au cours du temps

A partir de 2002, à l'exception de la décision n°2002-2725 AN (faux positif) et des deux décisions n°2022-5813/5814 AN, et n°2022-5760 AN (dysfonctionnements dans le vote électronique), toutes les annulations ont au moins un écart de voix entre les candidats 2 et 3 du tour 1 ou entre les candidats 1 et 2 du tour 2 inférieur ou égal à 279.

Pour la justice "predictive" : Enrichissement des données et annotation pour faire une analyse par griefs

- Très délicate tâche d'annotation.... L'échec de *Data Just*
- Début du travail d'annotation : quelle efficacité / coût (humain) ?

The screenshot displays the XML Author interface. The main window shows an XML document with the following content:

```
86 ;<br/> - SUR LES GRIEFS RELATIFS AU DÉROULEMENT DE LA CAMPAGNE ÉLECTORALE :<br/>
87 5. <Grief objet="4.Campagne_propagande_7_Lettres">Considérant que les lettres
88 mentionnées ci-dessus n'ont introduit dans la campagne électorale aucun
89 élément nouveau auquel M. CAHUZAC n'aurait pas eu le temps de répondre
90 ;</Grief><br/> 6. <Grief objet="4.Campagne_propagande_1_Affiches"
91 qualification="1.Grief_NC_2_Rejet_au_fond">Considérant que, s'il est soutenu
92 que M. MERLY se serait livré à un affichage en dehors des emplacements
93 officiels, en méconnaissance de l'article L. 51 du code électoral, il ne
94 ressort pas du procès-verbal de constat produit par le requérant que cet
95 affichage aurait revêtu un caractère massif ;</Grief><br/> 7. <Grief
96 objet="4.Campagne_propagande_16_Irrégularités_diverse"
97 qualification="1.Grief_NC_2_Rejet_au_fond">Considérant, enfin, que le grief
98 tiré d'une méconnaissance des dispositions de la loi du 6 janvier 1978,
99 relative à l'information sur les fichiers et aux libertés, n'est pas assorti
```

The interface includes a left sidebar with a tree view showing the document structure (e.g., META, TEXTE), a top menu bar, and a bottom status bar indicating a validation error: "Validation échouée. Erreurs: 1."

Traitement des données

Mise en forme de données déjà annotées selon plusieurs paramètres pour permettre des comparaisons de technologies et méthodologies

Deux modalités de TAL

1- Pipeline complète, sur Bigfoot

- Nécessite beaucoup de ressources, un accès et du temps de traitement

2- Chaîne de traitement allégée, tourne en local

- Dans une perspective de comparaison de technologies

Associées à plusieurs manières de représenter l'information

Séquençage des annotations sur la base des informations qu'elles contiennent, dans la dynamique exploratoire et comparative

8 processus parallèles, autant utiles pour le projet que dans l'aspect

→ méthodologie/reproductibilité

Des annotations plurielles

Un élément est annoté simultanément sur plusieurs aspects :

- L'objet du grief, ce qui concerne la demande
- la qualification, c'est à dire l'avis du juge

Chaque aspect comprend plusieurs informations (encadrés de couleur)

```
<Grief objet='6.Financement 5 Dépenses"  
qualification='1.Grief NC 2 Rejet au fond 2 non fondé">
```

→ Réduction des annotations pour opérer sur des "super-catégories", puis annotations des éléments dépendants sur lesdites "super-catégories"

Entraînement du modèle et annotation automatique

- Utilisation d'une pipeline orientée CPU pour qu'elle puisse tourner en local
- Pour chaque texte non-annoté :
 - Séparation des séquences à annoter par une série d'expressions régulières (“[Numéro]. “, puis par “;”)
 - Annotation automatique des séquences avec un seuil au choix de l'utilisateur
 - Reconstitution du fichier désormais annoté

Sur l'envoi des votes par correspondance aux bureaux de vote :

11. Considérant qu'il résulte de l'instruction que, contrairement aux dispositions de l'article R.87 du code électoral, 72 électeurs inscrits sur les listes électorales des communes de _____ ont voté par correspondance sans expédier leur vote sous pli recommandé ; que ces 72 votes sont nuls ; qu'il y a lieu, par application de la règle ci-dessus, d'opérer sur les suffrages obtenus par chacun des deux candidats la déduction de ces votes en fonction du résultat qu'ils ont obtenu dans chaque bureau de vote ; qu'en conséquence, 24 voix doivent être retranchées du nombre de suffrages recueillis par M. _____ et 48 voix doivent être déduites du nombre de suffrages dont M. _____ avait été crédité ;

12. Considérant que, contrairement aux allégations de la requête, il n'est pas établi que les votes par correspondance de 3 électeurs du 2e bureau de la ville de _____ aient été acceptés sans

Post-traitement des données annotées

- Récupération des annotations automatiques pour chaque fichier (procédure similaire au pré-traitement)
- formatage des données ainsi récupérées au format .csv en vue d'une vérification humaine

Vérification humaine	Remarque / état de l'annotation	Annotation alternative proposée	Fichier	Annotation	Texte
Oui	un considérant annoté sur 3		\AN\1963\annulation\CONSTEXT000017665174_annot.xml	7.Opérations	1. Considérant
Oui			\AN\1967\annulation\CONSTEXT000017665356_annot.xml	7.Opérations	4. Considérant
Oui			\AN\1967\annulation\CONSTEXT000017665356_annot.xml	7.Opérations	5. Considérant
Oui	considéran ts 3, 4, 5 ne sont pas annotés		\AN\1968\annulation\CONSTEXT000017665568_annot.xml	7.Opérations	2. Considérant
Oui	considérant 4 n'est pas annoté		\AN\1978\annulation\CONSTEXT000017665825_annot.xml	7.Opérations	3. Considérant
Oui	considéran ts 6, 8, 9, 10, 11, 12, 13, 16, 18 ne sont pas annotés		\AN\1978\annulation\CONSTEXT000017665826_annot.xml	7.Opérations	1. Considérant
Oui	considéran ts 6, 8, 9, 10, 11, 12, 13, 16, 18 ne sont pas annotés		\AN\1978\annulation\CONSTEXT000017665826_annot.xml	7.Opérations	3. Considérant
Oui	considéran ts 6, 8, 9, 10, 11, 12, 13, 16, 18 ne sont pas annotés		\AN\1978\annulation\CONSTEXT000017665826_annot.xml	7.Opérations	4. Considérant
Oui	considéran ts 6, 8, 9, 10, 11, 12, 13, 16, 18 ne sont pas annotés		\AN\1978\annulation\CONSTEXT000017665826_annot.xml	7.Opérations	5. Considérant
Oui	considéran ts 6, 8, 9, 10, 11, 12, 13, 16, 18 ne sont pas annotés		\AN\1978\annulation\CONSTEXT000017665826_annot.xml	7.Opérations	7. Considérant
Oui	considéran ts 6, 8, 9, 10, 11, 12, 13, 16, 18 ne sont pas annotés		\AN\1978\annulation\CONSTEXT000017665826_annot.xml	7.Opérations	14. Considérant
Oui	considéran ts 6, 8, 9, 10, 11, 12, 13, 16, 18 ne sont pas annotés		\AN\1978\annulation\CONSTEXT000017665826_annot.xml	7.Opérations	15. Considérant
Oui	considéran ts 5 et 7 ne sont pas annotés		\AN\1981\annulation\CONSTEXT000017665942_annot.xml	7.Opérations	4. Considérant
Non	considéran ts 3, 4 et 5 ne sont pas annotés	ne devrait être annoté (Rappel des faits, ici rappel des s	\AN\1981\annulation\CONSTEXT000017665946_annot.xml	7.Opérations	1. Considérant
Oui	considéran ts 3 et 4 ne sont pas annotés		\AN\1988\annulation\CONSTEXT000017667715_annot.xml	7.Opérations	2. Considérant
Non	considérant 7 n'est pas annoté	3. Candidature : éligibilité / inéligibilité	\AN\1988\annulation\CONSTEXT000017667716_annot.xml	4.Campagne	5. Considérant
Oui			\AN\1998\annulation\CONSTEXT000017667910_annot.xml	4.Campagne	1. Considérant
Oui			\AN\1998\annulation\CONSTEXT000017667910_annot.xml	4.Campagne	2. Considérant
Oui	considérant 4 n'est pas annoté		\AN\1999\annulation\CONSTEXT000017667949_annot.xml	7.Opérations	2. Considérant
Oui	considéran ts 2, 3 et 4 ne sont pas annotés		\AN\2002\annulation\CONSTEXT000017664506_annot.xml	6.Financement	: 1. Considérant
Oui	considéran ts 1 et 2 ne sont pas annotés		\AN\2002\annulation\CONSTEXT000017664522_annot.xml	6.Financement	4. Considérant
Oui	considéran ts 1 et 2 ne sont pas annotés		\AN\2002\annulation\CONSTEXT000017664522_annot.xml	6.Financement	5. Considérant

Une solution trouvée via dépôt sur Recherche Data Gouv

- Bligny, Caroline; Rambaud, Romain, 2024, “Cross-referenced data on electoral disputes and French legislative election results”, <https://doi.org/10.57745/0BI3LB>, Recherche Data Gouv, V1.

2 "dépôts" :

- Le dépôt du Dataset jade sur recherche.data.gouv.fr, pour lequel les meta données sont visibles mais pas les données, tant que les questions rgpd ne sont pas claires.-
 - Le dépôt du code de construction et d'exploitation de la BDD jade sur gitlab (qui contient aussi le code de Louis), avec une licence ouverte mais accès restreint aussi pour la même raison
-
- Accès pour l’instant **restreint** du fait des questions RGPD & CNIL
 - Par la suite ?

Avantages et limites de la collaboration JADE / GATES

- **Enorme avantage pour JADE** : l'aide de GATES pendant 6 mois a permis au projet JADE de commencer à expérimenter le TAL pour aller plus loin (justice prédictive)
- **Limites** :
 - Un mi-temps pendant 6 mois n'a pas été suffisant pour avoir de bons résultats, il a fallu plus de temps
 - Le projet JADE a proposé qu'un article soit écrit pour conserver les résultats de cette expérimentation, mais il faut du temps laissé à l'ingénieur pour l'écrire, et il y a des débats concernant l'état de l'art (pluridisciplinarité)
 - La pérennité de JADE dépend aujourd'hui de la réussite à une ANR aux très faibles taux de réussite (réponse phase 1 en février)
 - Si pas d'article et/ou pas d'ANR, tout le travail effectué va se perdre définitivement

Pour en savoir plus

Blog du droit électoral :



blogdudroitelectoral.fr/justice-algorithmique-des-elections-jade/

Blog du pool d'ingénieurs GATES



gates-data-shs.gricad-pages.univ-grenoble-alpes.fr/blog/

Publications

- 1) Rambaud R., « Les "lois" de l'écart de voix », AJDA, 2020, p. 1596.
- 2) Rambaud R., Hafsaoui A., Bligny C., « Une justice algorithmique pour les élections ? », AJDA, 2023, p. 1323.
- 3) Rambaud R., C. Bligny, A. Hafsaoui, S. Cottin, « Justice algorithmique des élections (JADE): une nouvelle base de données du contentieux des élections parlementaires devant le Conseil constitutionnel », Jurimétrie, décembre 2024.
- 4) Rambaud R., C. Bligny, F. Letué, M.-J. Martinez, S. Cottin, J.-P. Camby, G. Prunier, D. Girard, A. Hafsaoui, K. Deschamps, "Projet Justice algorithmique des élections (JADE) : une analyse statistique de la jurisprudence du Conseil constitutionnel relative aux élections législatives (4 oct. 1958 – 1er avril 2024). Partie 1/2 – La structuration générale du contentieux AN, Revue française de droit constitutionnel, n°140, janvier 2025.
- 5) Rambaud R., C. Bligny, F. Letué, M.-J. Martinez, S. Cottin, J.-P. Camby, G. Prunier, D. Girard, A. Hafsaoui, K. Deschamps, "Projet Justice algorithmique des élections (JADE) : une analyse statistique de la jurisprudence du Conseil constitutionnel relative aux élections législatives (4 oct. 1958 – 1er avril 2024), Partie 2/2 – Sincérité du scrutin et critère de l'écart de voix", n°141, mars 2025.