TRAITEMENT ET ANALYSE DES DÉBATS PARLEMENTAIRES À LA CHAMBRE DES DÉPUTÉS (1881-1899)

Problèmes, défis et solutions

Aurélien Pellet[1] et Marie Puren[2]

10 novembre 2022 - Res(t)ituer les adresses des almanachs et annuaires commerciaux parisiens du XIXe siècle. Un corpus de localisations urbaines à grande échelle. 2e Journée de l'atelier SoDUCo-BnF

[1]Epitech [2]LRE, EPITA

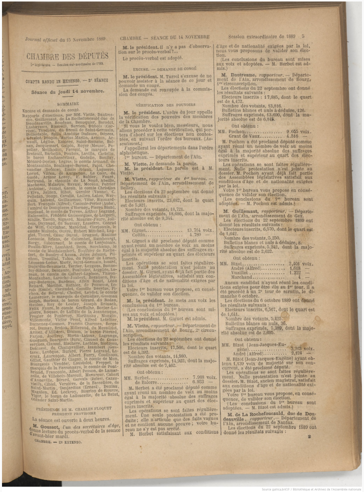


Figure – Séance parlementaire du 14 novembre 1889

- Projet AGODA : **A**nalyse sémantique et **G**raphes relationnels pour l'**O**uverture et l'étude des **D**ébats à l'**A**ssemblée nationale
- Débats à la Chambre des députés (chambre basse du parlement) transcrits en détail dans le Journal officiel de la République française. Débats parlementaires (1881-1940)
- Disponible en ligne via Gallica (bibliothèque numérique de la Bibliothèque nationale de France)
- 15 législatures entre 1881 et 1940 / 10-12.000 images par législature

OBJECTIFS

- Créer une plateforme de consultation
- Produire des données textuelles structurées et sémantiquement enrichies à partir de ces débats numérisés
- Contribuer à la conception d'un workflow adapté à la préparation, à la publication et à l'analyse de grands corpus de documents historiques

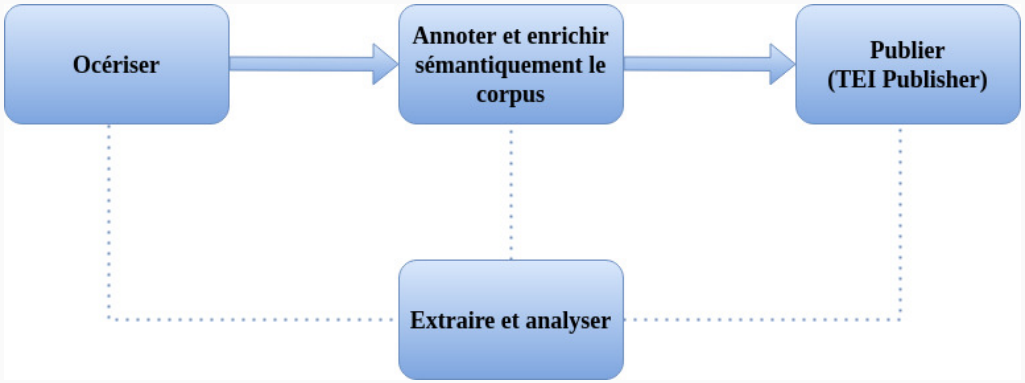


Figure – Les étapes de la chaîne de traitement

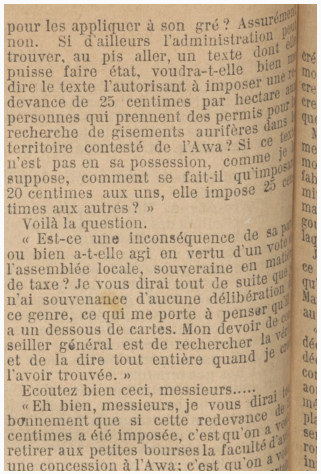
OCÉRISER LES DÉBATS

LE CAS DES DÉBATS PARLEMENTAIRES

- Récupération des textes océrisés via API Document de Gallica => qualité inégale de l'OCR
- Erreurs dues à :
  - qualité du document : tâches et surimpression
  - la **courbure de la page** au niveau de la reliure

Figure – Evaluation de la qualité de l'OCR fourni par Gallica

EFFET DE LA COURBURE SUR LES RÉSULTATS DE L'OCR



pour les appliquer à son gré ? Assurém
non. Si d'ailleurs l'administration pe
trouver, au pis aller, un texte dont é
puisse faire état, voudra-t-elle bien
dire le texte l'autorisant à imposer une
devance de 25 centimes par hectare a
personnes qui prennent des permis pour
recherche de gisements aurifères dans
territoire contesté de l'Awa ? Si ce te
n'est pas en sa possession, comme je
suppose, comment se fait-il qu'impos
20 centimes aux uns, elle impose 25 ce
times aux autres ? »
Voilà la question.
« Est-ce une inconséquence de sa p
ou bien a-t-elle agi en vertu d'un vot
l'assemblée locale, souveraine en mati
de taxe ? Je vous dirai tout de suite qu
n'ai souvenance d'aucune délibération
ce genre, ce qui me porte à penser qu
a un dessous de cartes. Mon devoir de c
seiller général est de rechercher la vé
et de la dire tout entière quand je
l'avoir trouvée. »
Ecoutez bien ceci, messieurs.....
« Eh bien, messieurs, je vous dirai
bonnement que si cette redevance de
centimes a été imposée, c'est qu'on a
retirer aux petites bourses la faculté d'
une concession à l'Awa; c'est qu'on a

Figure – 20 octobre 1890 (p.1718)

pour les appliquer à son gré ? Assure.

non. Si d'ailleurs l'administration fj, trouver, au pis aller, un texte (1011t eu puisse faire état, voudra-t-elle dire le texte l'autorisant à imposer u devance de 25 centimes par hectar , personnes qui prennent des permis Pour1 recherche de gisements aurifères a teSI territoire contesté de l'Awa ? Si ce tes

n'est pas en sa possession, cornas le suppose, comment se fait-il qulIWjs3j 20 centimes aux uns, elle impose 20 cet times aux autres ? » , Voilà la question. é

« Est-ce une inconséquence de sa te ou bien a-t-elle agi en vertu d'un v l'assemblée locale, souveraine en OlqallC de taxe? Je vous dirai tout de suite Il n'ai souvenance d'aucune délibératl°ll'V ce genre, ce qui me porte à penser a un dessous de cartes. Mon devoir seiller général est de rechercher la cJ'!" et de la dire tout entière quand Je l'avoir trouvée. »

Ecoutez bien ceci, messieurs. l «Eh bien, messieurs, je vous dIra de bonnement que si cette redevance J centimes a été imposée, c'est qu'on a, retirer aux petites bourses la faculté" une concession à l'Awa; c'est qu'on a r¡;:i\

Figure – Résultat de l'OCR

Décision de ré-océriser le corpus

TRAITER LES IMAGES

Améliorer la qualité de l'image avec une méthode de "dewarping" => résultats peu probants

- Gérer la courbure des pages avec le dewarping ?
- Utiliser des outils plus avancés ?

(a) Image d'origine

(b) Image "dewarpée"

Figure – Dewarping : pas adapté à nos documents

OUTIL DE NETTOYAGE - ANR SODUCO

Année 1881 DÉBATS ET DOCUMENTS PARLEMENTAIRES 25 Février [CHAMBRE] 345

CHAMBRE DES DÉPUTÉS
Session ordinaire de 1881

COMPTE RENDU IN EXTENSO. — 23e SÉANCE

Séance du jeudi 24 février 1881

SOMMAIRE

Procès-verbal : M. Devès.
Excuse.
Dépôt, par M. Marquiset, du rapport de la commission chargée d'examiner le projet de loi ayant pour objet de faire déclarer d'utilité publique la construction d'un canal de navigation dans la vallée de la Chiers, entre Longwy et la Meuse, à Mouzon.
Dépôt, par M. Morel, du rapport de la 25e commission d'intérêt local, sur le projet de loi tendant à autoriser le département de la Haute-Marne à contracter un emprunt pour l'amélioration de l'asile des aliénés de Saint-Dizier.
Dépôt, par M. Sarrien, du rapport de la commission chargée d'examiner le projet de loi ayant pour objet de faire déclarer d'utilité publique l'amélioration du canal du centre.
Suite de la 1re délibération sur le projet de loi, adopté par le Sénat, concernant l'administration de l'armée : M. le comte de Roys, rapporteur. — Ajournement à samedi.
Adoption en 2e délibération du projet de loi ayant pour objet l'établissement d'un canal de Dombasle à Saint-Dié.
Discussion des conclusions du rapport de la 2e commission des pétitions sur la pétition du sieur Jean Dupuis, citoyen français, demeurant à Han-Kow (Chine) : M. Emile Bouchet, rapporteur. Adoption.
1re délibération sur les propositions de loi : 1° de M. Germain Casse et de plusieurs de ses collègues, tendant à régler les rapports entre les compagnies de chemins de fer et leurs agents commissionnés ; 2° de M. le baron de Janzé, relative à la réglementation des compagnies de chemins de fer : MM. René Brice, le baron de Janzé.
Règlement de l'ordre du jour : MM. Latrade, Sarrette, de Gasté, Bardoux.
Reprise de la discussion des propositions tendant à régler les rapports entre les compagnies de chemins de fer et leurs agents commissionnés : M. Trarieux.
Présentation, par M. le ministre de l'intérieur et des cultes, d'un projet de loi tendant à autoriser le département des Côtes-du-Nord à contracter deux emprunts pour les travaux des chemins vicinaux ordinaires.
Demande, par M. le ministre de la marine, de discussion immédiate du projet de loi, adopté par la Chambre des députés, adopté avec modifications par le Sénat, ayant pour objet d'ouvrir au ministre de la marine et des colonies un crédit de 8,352,751 fr., sur l'exercice 1881, pour les chemins de fer du Haut-Sénégal. Adoption. = Article 1er : M. Des Rotours. Adoption. = Adoption de l'article 2. = Adoption, au scrutin, de l'ensemble du projet de loi.
Règlement de l'ordre du jour : MM. le président, de Gasté.
Présentation par M. le sous-secrétaire d'Etat des travaux publics :
1° D'un projet de loi ayant pour objet : 1° L'incorporation dans le réseau d'intérêt général du chemin de fer d'intérêt local de Mézidon à Dives ; 2° l'approbation d'une convention passée avec la compagnie des chemins de fer de la vallée d'Auge pour le rachat de ladite ligne ;
2° D'un projet de loi ayant pour objet : 1° De déclarer d'utilité publique l'établissement, dans le département des Landes, de plusieurs chemins de fer d'intérêt local et d'en autoriser l'exécution ; 2° d'approuver un traité passé entre la compagnie du Midi et les concessionnaires de ces chemins.
Dépôt par M. Cavalié d'un rapport de la 25e commission d'intérêt local sur le projet de loi tendant à autoriser le département de la Haute-Saône à contracter un emprunt pour les travaux des chemins vicinaux.

PRÉSIDENCE DE M. GAMBETTA

La séance est ouverte à deux heures.
M. Frémfact, *l'un des secrétaires*, donne lecture du procès-verbal de la séance de mardi dernier.
M. le président. La parole est à M. Devès sur le procès-verbal.
M. Paul Devès. Messieurs, MM. Tassin, de Sonnier et Deniau auraient voté contre la proposition d'enquête faite par M. Lorgé s'ils eussent assisté à la séance de mardi. Je le déclare en leur nom.
M. le président. Il n'y a pas d'autre observation ?...
Le procès-verbal est adopté.

EXCUSE

M. le président. M. Dremel s'excuse de ne pouvoir assister à la séance de ce jour.

DÉPÔT DE RAPPORTS

M. le président. Je donne la parole à ceux de MM. les députés qui ont des rapports à déposer.
M. Sarrien. Au nom de la commission des voies navigables, j'ai l'honneur de déposer sur le bureau de la Chambre un rapport sur le projet de loi ayant pour objet de faire déclarer d'utilité publique l'amélioration du canal du Centre.
M. Marquiset. J'ai l'honneur de déposer sur le bureau de la Chambre, au nom de la commission des voies navigables, un rapport sur le projet de loi ayant pour objet de faire déclarer d'utilité publique la construction d'un canal de navigation dans la vallée de la Chiers, entre Longwy et la Meuse, à Mouzon.
M. Morel. Au nom de la 25e commission d'intérêt local, j'ai l'honneur de déposer sur le bureau de la Chambre un rapport sur le projet de loi tendant à autoriser le département de la Haute-Marne à contracter un emprunt pour l'amélioration de l'asile des aliénés de Saint-Dizier.
M. le président. Ces rapports seront imprimés et distribués.

AJOURNEMENT DE LA SUITE DE LA DISCUSSION DU PROJET DE LOI SUR L'ADMINISTRATION DE L'ARMÉE.

M. le président. L'ordre du jour appelle la suite de la première délibération sur le projet de loi, adopté par le Sénat, concernant l'administration de l'armée.
M. le rapporteur a la parole.
M. le comte de Roys, *rapporteur*. Messieurs, M. le ministre de la guerre étant appelé aujourd'hui au Sénat par une interpellation, la commission vous prie de vouloir bien ajourner à la séance de samedi la suite de la discussion sur l'administration de l'armée, en lui maintenant son rang en tête de l'ordre du jour de cette séance.
M. le président. Il n'y a pas d'opposition ?...
L'ajournement à samedi est prononcé.

DISCUSSION DU PROJET DE LOI CONCERNANT LE CANAL DE DOMBASLE A SAINT DIÉ

M. le président. L'ordre du jour appelle la deuxième délibération sur le projet de loi ayant pour objet l'établissement d'un canal de Dombasle à Saint-Dié.
M. le président. Le rapport a été imprimé et distribué, et la discussion est inscrite au feuilleton depuis deux mois.
M. de Larochefoucauld, duc de Bisaccia. Avant de voter, je demande des explications.
M. le président. La commission conclut au renvoi à M. le ministre de la marine.
ponts et chaussées dans ses délibérations des 7 novembre 1878 et 29 janvier 1880. »
(L'article 1er est mis aux voix et adopté.)
« Art. 2. — Un péage dont le produit est garanti, comme il est dit à l'article 3, par les départements des Vosges et de Meurthe-et-Moselle, sera établi sur la voie navigable, au fur et à mesure de la mise en exploitation de ses diverses sections.
« Ce péage se composera, par tonne, d'une constante de 0.30 et d'un élément variable de 0,0005 par kilomètre.
« Il prendra fin vingt ans après l'ouverture du canal sur toute sa longueur. » — (Adopté.)
« Art. 3. — Le ministre des travaux publics est autorisé à accepter, au nom de l'Etat :
« 1° Les subventions souscrites par les industriels intéressés à titre de libération du péage ;
« 2° L'offre faite par les deux départements de Meurthe-et-Moselle et des Vosges, ainsi qu'il résulte des délibérations des conseils généraux en date des 22 et 23 août 1879, de garantir solidairement et à parts égales les subventions précitées ainsi qu'un produit annuel de 140,000 francs pour le péage à partir de la mise en exploitation du canal sur toute sa longueur, sauf déduction, au taux de 4 1/2 p. 100, de l'annuité d'amortissement en vingt ans de ces subventions ;
« 3° Les offres de concours des villes de Lunéville, Baccarat, Raon l'Etape, Saint-Dié et Fraize montant ensemble à la somme de 5 10,000 francs. » — (Adopté.)
« Art. 4. — La somme de 20 millions de francs à laquelle sont évalués les travaux sera imputée sur les ressources extraordinaires inscrites au budget de chaque année.
« Viendra en déduction de la dépense à la charge du Trésor le montant des subventions mentionnées à l'article précédent. » — (Adopté.)
(L'ensemble du projet de loi est mis aux voix et adopté.)

DISCUSSION D'UN RAPPORT DE LA 2e COMMISSION DES PÉTITIONS

M. le président. L'ordre du jour appelle la discussion des conclusions du rapport fait au nom de la 2e commission des pétitions sur la pétition du sieur Jean Dupuis, citoyen français, demeurant à Han-Kow (Chine).
La parole est à M. le rapporteur.
M. Emile Bouchet, *rapporteur*. Messieurs, sur cette question qui aurait comporté de longs développements, je n'aurai que très-peu de mots à prononcer : je demande simplement à la Chambre de vouloir bien accueillir les conclusions de la commission, tendant au renvoi de la pétition au ministre de la marine, conclusions acceptées par M. le ministre de la marine et par le Gouvernement.
M. de Mahy. Très-bien ! très-bien !
M. La Vieille. Quelles sont ces conclusions ?
M. le président. La commission conclut au renvoi de la pétition au ministre de la marine ; le renvoi est accepté par le Gouvernement.
M. le rapporteur. Les conclusions sont celles de toutes les commissions de pétitions qui, après avoir examiné les faits, concluent, quand elles les trouvent dignes d'une attention particulière, au renvoi de la pétition au ministre compétent. C'est ce qu'a fait la deuxième commission des pétitions.
M. de Larochefoucauld, duc de Bisaccia. Quelle est cette pétition ?
M. le président. Le rapport a été imprimé et distribué, et la discussion est inscrite au feuilleton depuis deux mois.
M. de Larochefoucauld, duc de Bisaccia. Avant de voter, je demande des explications.
M. le président. La commission conclut au renvoi à M. le ministre de la marine.

(a) Image d'origine

Année 1881 DÉBATS ET DOCUMENTS PARLEMENTAIRES 25 Février [CHAMBRE] 345

CHAMBRE DES DÉPUTÉS
Session ordinaire de 1881

COMPTE RENDU IN EXTENSO. — 23e SÉANCE

Séance du jeudi 24 février 1881

SOMMAIRE

Procès-verbal : M. Devès.
Excuse.
Dépôt, par M. Marquiset, du rapport de la commission chargée d'examiner le projet de loi ayant pour objet de faire déclarer d'utilité publique la construction d'un canal de navigation dans la vallée de la Chiers, entre Longwy et la Meuse, à Mouzon.
Dépôt, par M. Morel, du rapport de la 25e commission d'intérêt local, sur le projet de loi tendant à autoriser le département de la Haute-Marne à contracter un emprunt pour l'amélioration de l'asile des aliénés de Saint-Dizier.
Dépôt, par M. Sarrien, du rapport de la commission chargée d'examiner le projet de loi ayant pour objet de faire déclarer d'utilité publique l'amélioration du canal du centre.
Suite de la 1re délibération sur le projet de loi, adopté par le Sénat, concernant l'administration de l'armée : M. le comte de Roys, rapporteur. — Ajournement à samedi.
Adoption en 2e délibération du projet de loi ayant pour objet l'établissement d'un canal de Dombasle à Saint-Dié.
Discussion des conclusions du rapport de la 2e commission des pétitions sur la pétition du sieur Jean Dupuis, citoyen français, demeurant à Han-Kow (Chine) : M. Emile Bouchet, rapporteur. Adoption.
1re délibération sur les propositions de loi : 1° de M. Germain Casse et de plusieurs de ses collègues, tendant à régler les rapports entre les compagnies de chemins de fer et leurs agents commissionnés ; 2° de M. le baron de Janzé, relative à la réglementation des compagnies de chemins de fer : MM. René Brice, le baron de Janzé.
Règlement de l'ordre du jour : MM. Latrade, Sarrette, de Gasté, Bardoux.
Reprise de la discussion des propositions tendant à régler les rapports entre les compagnies de chemins de fer et leurs agents commissionnés : M. Trarieux.
Présentation, par M. le ministre de l'intérieur et des cultes, d'un projet de loi tendant à autoriser le département des Côtes-du-Nord à contracter deux emprunts pour les travaux des chemins vicinaux ordinaires.
Demande, par M. le ministre de la marine, de discussion immédiate du projet de loi, adopté par la Chambre des députés, adopté avec modifications par le Sénat, ayant pour objet d'ouvrir au ministre de la marine et des colonies un crédit de 8,352,751 fr., sur l'exercice 1881, pour les chemins de fer du Haut-Sénégal. Adoption. = Article 1er : M. Des Rotours. Adoption. = Adoption de l'article 2. = Adoption, au scrutin, de l'ensemble du projet de loi.
Règlement de l'ordre du jour : MM. le président, de Gasté.
Présentation par M. le sous-secrétaire d'Etat des travaux publics :
1° D'un projet de loi ayant pour objet : 1° L'incorporation dans le réseau d'intérêt général du chemin de fer d'intérêt local de Mézidon à Dives ; 2° l'approbation d'une convention passée avec la compagnie des chemins de fer de la vallée d'Auge pour le rachat de ladite ligne ;
2° D'un projet de loi ayant pour objet : 1° De déclarer d'utilité publique l'établissement, dans le département des Landes, de plusieurs chemins de fer d'intérêt local et d'en autoriser l'exécution ; 2° d'approuver un traité passé entre la compagnie du Midi et les concessionnaires de ces chemins.
Dépôt par M. Cavalié d'un rapport de la 25e commission d'intérêt local sur le projet de loi tendant à autoriser le département de la Haute-Saône à contracter un emprunt pour les travaux des chemins vicinaux.

PRÉSIDENCE DE M. GAMBETTA

La séance est ouverte à deux heures.
M. Frémfact, *l'un des secrétaires*, donne lecture du procès-verbal de la séance de mardi dernier.
M. le président. La parole est à M. Devès sur le procès-verbal.
M. Paul Devès. Messieurs, MM. Tassin, de Sonnier et Deniau auraient voté contre la proposition d'enquête faite par M. Lorgé s'ils eussent assisté à la séance de mardi. Je le déclare en leur nom.
M. le président. Il n'y a pas d'autre observation ?...
Le procès-verbal est adopté.

EXCUSE

M. le président. M. Dremel s'excuse de ne pouvoir assister à la séance de ce jour.

DÉPÔT DE RAPPORTS

M. le président. Je donne la parole à ceux de MM. les députés qui ont des rapports à déposer.
M. Sarrien. Au nom de la commission des voies navigables, j'ai l'honneur de déposer sur le bureau de la Chambre un rapport sur le projet de loi ayant pour objet de faire déclarer d'utilité publique l'amélioration du canal du Centre.
M. Marquiset. J'ai l'honneur de déposer sur le bureau de la Chambre, au nom de la commission des voies navigables, un rapport sur le projet de loi ayant pour objet de faire déclarer d'utilité publique la construction d'un canal de navigation dans la vallée de la Chiers, entre Longwy et la Meuse, à Mouzon.
M. Morel. Au nom de la 25e commission d'intérêt local, j'ai l'honneur de déposer sur le bureau de la Chambre un rapport sur le projet de loi tendant à autoriser le département de la Haute-Marne à contracter un emprunt pour l'amélioration de l'asile des aliénés de Saint-Dizier.
M. le président. Ces rapports seront imprimés et distribués.

AJOURNEMENT DE LA SUITE DE LA DISCUSSION DU PROJET DE LOI SUR L'ADMINISTRATION DE L'ARMÉE.

M. le président. L'ordre du jour appelle la suite de la première délibération sur le projet de loi, adopté par le Sénat, concernant l'administration de l'armée.
M. le rapporteur a la parole.
M. le comte de Roys, *rapporteur*. Messieurs, M. le ministre de la guerre étant appelé aujourd'hui au Sénat par une interpellation, la commission vous prie de vouloir bien ajourner à la séance de samedi la suite de la discussion sur l'administration de l'armée, en lui maintenant son rang en tête de l'ordre du jour de cette séance.
M. le président. Il n'y a pas d'opposition ?...
L'ajournement à samedi est prononcé.

DISCUSSION DU PROJET DE LOI CONCERNANT LE CANAL DE DOMBASLE A SAINT DIÉ

M. le président. L'ordre du jour appelle la deuxième délibération sur le projet de loi ayant pour objet l'établissement d'un canal de Dombasle à Saint-Dié.
M. le président. Le rapport a été imprimé et distribué, et la discussion est inscrite au feuilleton depuis deux mois.
M. de Larochefoucauld, duc de Bisaccia. Avant de voter, je demande des explications.
M. le président. La commission conclut au renvoi à M. le ministre de la marine.
ponts et chaussées dans ses délibérations des 7 novembre 1878 et 29 janvier 1880. »
(L'article 1er est mis aux voix et adopté.)
« Art. 2. — Un péage dont le produit est garanti, comme il est dit à l'article 3, par les départements des Vosges et de Meurthe-et-Moselle, sera établi sur la voie navigable, au fur et à mesure de la mise en exploitation de ses diverses sections.
« Ce péage se composera, par tonne, d'une constante de 0.30 et d'un élément variable de 0,0005 par kilomètre.
« Il prendra fin vingt ans après l'ouverture du canal sur toute sa longueur. » — (Adopté.)
« Art. 3. — Le ministre des travaux publics est autorisé à accepter, au nom de l'Etat :
« 1° Les subventions souscrites par les industriels intéressés à titre de libération du péage ;
« 2° L'offre faite par les deux départements de Meurthe-et-Moselle et des Vosges, ainsi qu'il résulte des délibérations des conseils généraux en date des 22 et 23 août 1879, de garantir solidairement et à parts égales les subventions précitées ainsi qu'un produit annuel de 140,000 francs pour le péage à partir de la mise en exploitation du canal sur toute sa longueur, sauf déduction, au taux de 4 1/2 p. 100, de l'annuité d'amortissement en vingt ans de ces subventions ;
« 3° Les offres de concours des villes de Lunéville, Baccarat, Raon l'Etape, Saint-Dié et Fraize montant ensemble à la somme de 5 10,000 francs. » — (Adopté.)
« Art. 4. — La somme de 20 millions de francs à laquelle sont évalués les travaux sera imputée sur les ressources extraordinaires inscrites au budget de chaque année.
« Viendra en déduction de la dépense à la charge du Trésor le montant des subventions mentionnées à l'article précédent. » — (Adopté.)
(L'ensemble du projet de loi est mis aux voix et adopté.)

DISCUSSION D'UN RAPPORT DE LA 2e COMMISSION DES PÉTITIONS

M. le président. L'ordre du jour appelle la discussion des conclusions du rapport fait au nom de la 2e commission des pétitions sur la pétition du sieur Jean Dupuis, citoyen français, demeurant à Han-Kow (Chine).
La parole est à M. le rapporteur.
M. Emile Bouchet, *rapporteur*. Messieurs, sur cette question qui aurait comporté de longs développements, je n'aurai que très-peu de mots à prononcer : je demande simplement à la Chambre de vouloir bien accueillir les conclusions de la commission, tendant au renvoi de la pétition au ministre de la marine, conclusions acceptées par M. le ministre de la marine et par le Gouvernement.
M. de Mahy. Très-bien ! très-bien !
M. La Vieille. Quelles sont ces conclusions ?
M. le président. La commission conclut au renvoi de la pétition au ministre de la marine ; le renvoi est accepté par le Gouvernement.
M. le rapporteur. Les conclusions sont celles de toutes les commissions de pétitions qui, après avoir examiné les faits, concluent, quand elles les trouvent dignes d'une attention particulière, au renvoi de la pétition au ministre compétent. C'est ce qu'a fait la deuxième commission des pétitions.
M. de Larochefoucauld, duc de Bisaccia. Quelle est cette pétition ?
M. le président. Le rapport a été imprimé et distribué, et la discussion est inscrite au feuilleton depuis deux mois.
M. de Larochefoucauld, duc de Bisaccia. Avant de voter, je demande des explications.
M. le président. La commission conclut au renvoi à M. le ministre de la marine.

(b) Image nettoyée

Figure – Exemple d'image nettoyée

OUTIL OCR ET NER - ANR SODUCO

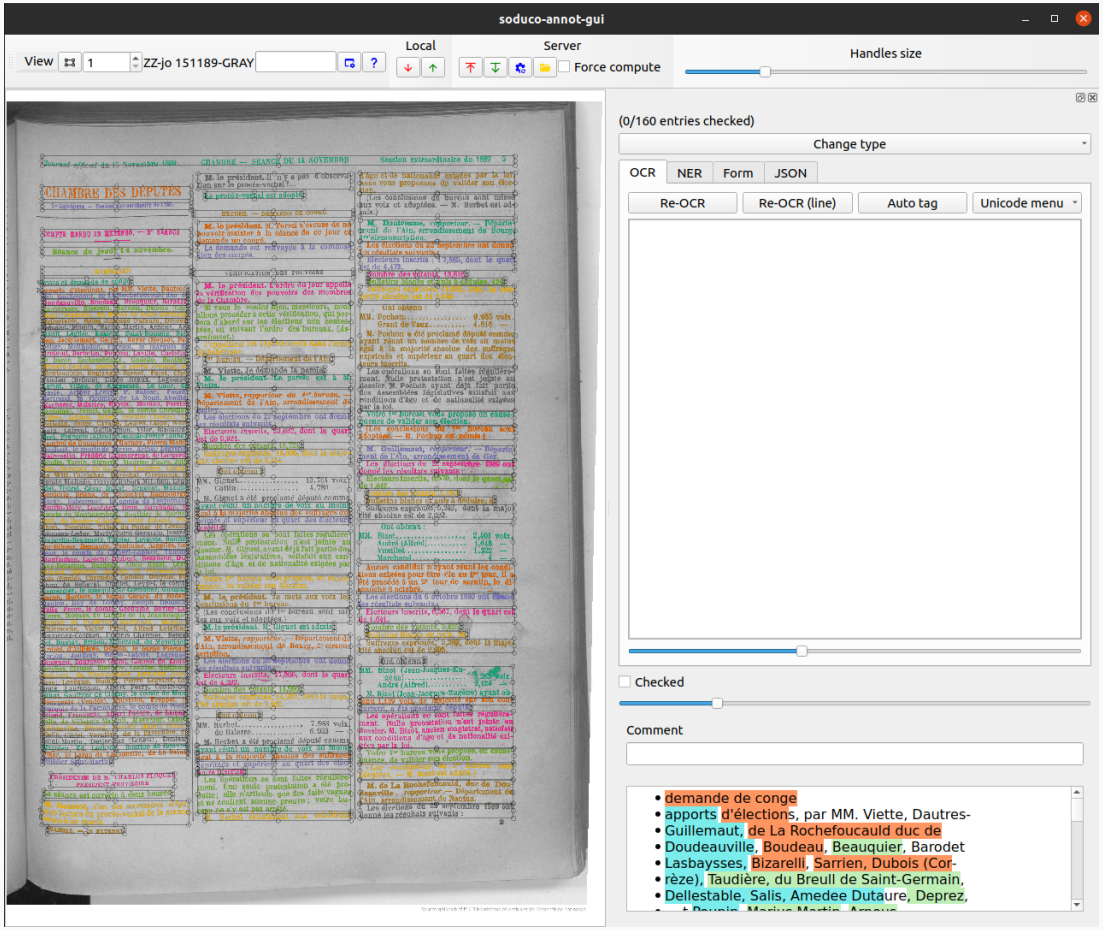
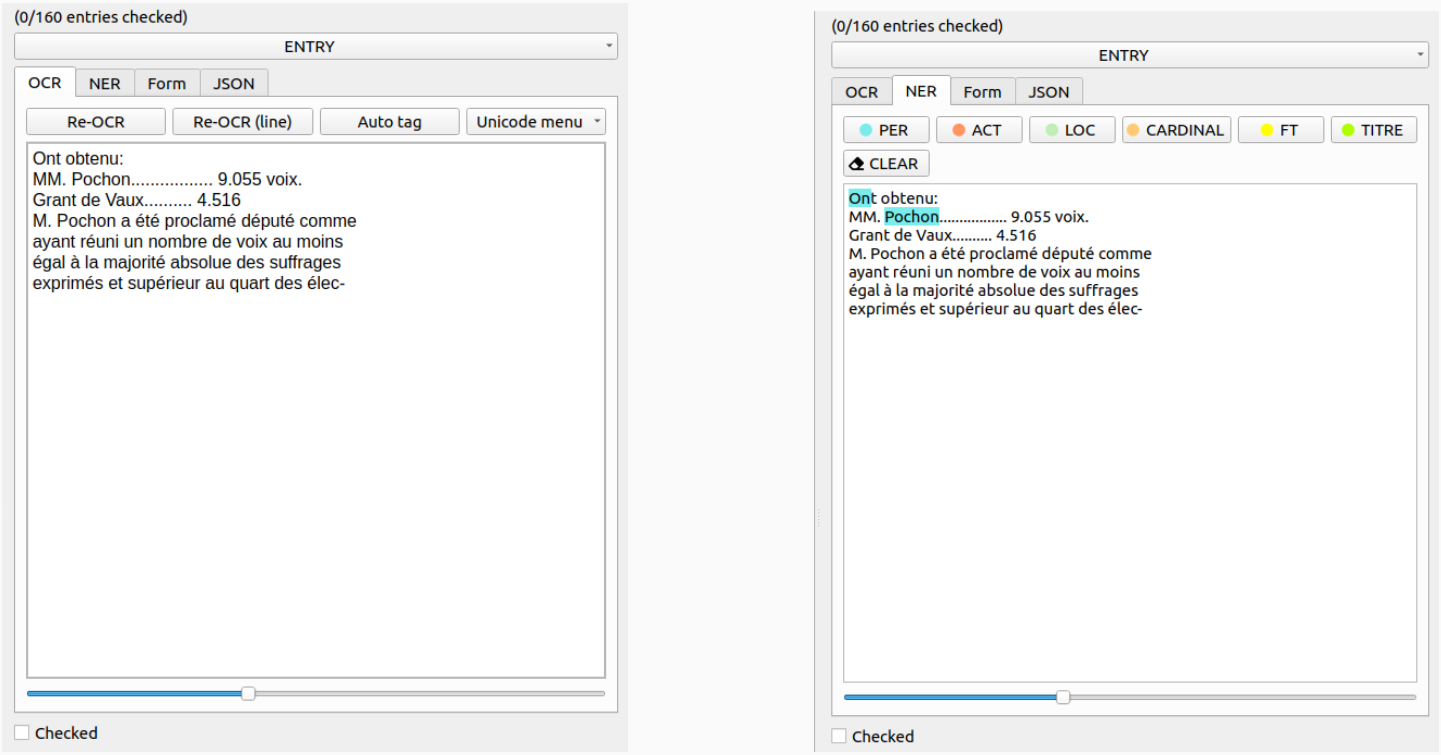


Figure – Démonstration de l'outil sur une page de débat parlementaire

OUTIL OCR ET NER - ANR SODUCO



(a) OCR **(b)** NER

Figure – Zones d'OCR et de NER

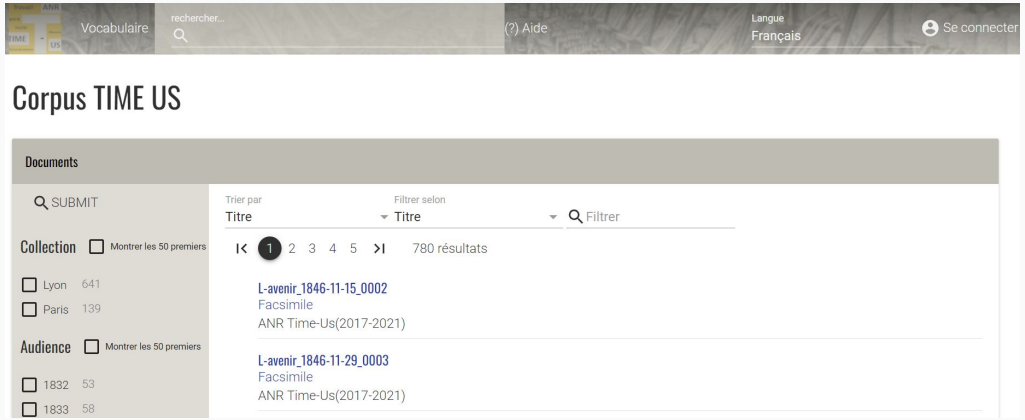
AMÉLIORER LA COURBURE

- Problèmes de reconnaissance de la structure hiérarchique à cause de la courbure
- Objectif n°1 : faire disparatre au maximum cette courbure
- Développement et traitement d'images en C++

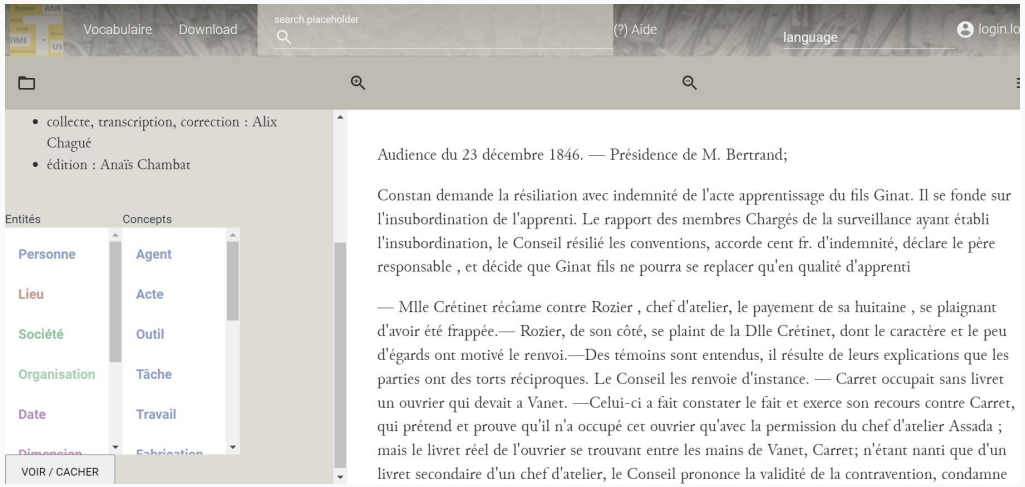
Demande de financement dans le cadre de Biblissima +

Annoter les débats en XML-TEI

PUBLIER UN CORPUS ÉDITORIALISÉ



(a) Interface de recherche



(b) Affichage d'un document

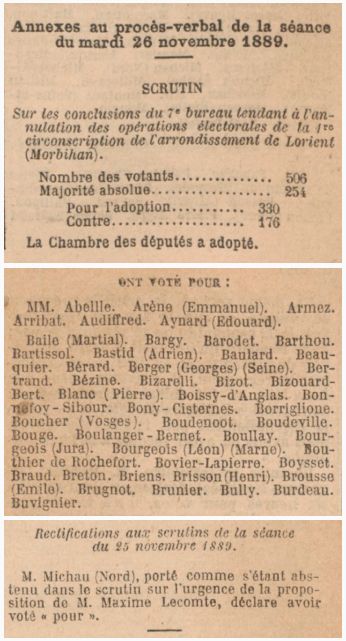
Figure – Corpus du projet ANR TIME-US publié avec TEI Publisher

CORPS D'UN DÉBAT ENCODÉ EN XML-TEI

```
        <div type="part" corresp="#rapportelections">
           <head>SUITE DE LA VÉRIFICATION DES POUVOIRS</head>
           <u who="#pers_ID" xml:id="CR_1889-11-26_u23" ana="#chair">
              <seg xml:id="CR_1889-11-26_u23.1"><persName ref="#pers_ID">M. le <roleName ref="#pers_ID">président</roleName></persName>. L'ordre du jour
appelle la suite de la vérification des pouvoirs.</seg>
              <seg xml:id="CR_1889-11-26_u23.2"><persName ref="#pers_ID">M. Reybert</persName> a la parole pour donner lecture d'un rapport sur une
élection non contestée.</seg>
           </u>

           <u who="#pers_ID" xml:id="CR_1889-11-26_u24" ana="#rapporteur">
              <!-- Lecture d'un rapport -->
              <quote>
                 <seg xml:id="CR_1889-11-26_u24.1"><persName ref="#pers_ID">M. Reybert, <roleName ref="#pers_ID">rapporteur</roleName></persName>.
—<placeName ref="#lieu_ID">Département de la Corrèze, arrondissement de Tulle, <num>1re</num> circonscription</placeName>.</seg>
                 <seg xml:id="CR_1889-11-26_u24.2">Les élections du <date when="1889-09-22">22 septembre</date> ont donné les résultats suivants :</seg>
                 <seg xml:id="CR_1889-11-26_u24.3">Electeurs inscrits, <num>17,950</num>, dont le quart est de <num>4,263</num>.</seg>
                 <seg xml:id="CR_1889-11-26_u24.4">Nombre des votants, <num>12,322</num>.</seg>
                 <seg xml:id="CR_1889-11-26_u24.5">Bulletins blancs et nuls, à déduire, <num>133</num>.</seg>
                 <seg xml:id="CR_1889-11-26_u24.6">Suffrages exprimés, <num>12,189</num>.</seg>
                 <seg xml:id="CR_1889-11-26_u24.7">Ont obtenu : MM. <persName ref="#pers_ID">Borie, <roleName ref="#pers_ID">député
sortant</roleName></persName>... <num>7.508</num> voix</seg>
                 <seg xml:id="CR_1889-11-26_u24.8"><persName ref="#pers_ID">Vachal, <roleName ref="#pers_ID">ancien député</roleName></persName>...
<num>4.748</num></seg>
                 <seg xml:id="CR_1889-11-26_u24.9"><persName ref="#pers_ID">M. Borie</persName> a été proclamé député comme ayant réuni un nombre de voix
au moins égal à la majorité absolue des suffrages</seg>
              </quote>
           </u>

           <floatingText><body><div><pb n="176"/></div></body></floatingText>

           <!-- [...] -->
        </div>
```

Figure – Encodage - Séance parlementaire du 26 novembre 1889 (extrait)

ANNEXES D'UN DÉBAT ENCODÉES EN XML-TEI



Annexes au procès-verbal de la séance du mardi 26 novembre 1889.

SCRUTIN

Sur les conclusions du 7e bureau tendant à l'annulation des opérations électorales de la 1re circonscription de l'arrondissement de Lorient (Morbihan).

Nombre des votants..... 506
Majorité absolue..... 254
Pour l'adoption..... 330
Contre..... 176

La Chambre des députés a adopté.

ONT VOTÉ POUR :

MM. Abeille. Arène (Emmanuel). Armez. Arribat. Audiffred. Aynard (Edouard). Baile (Martial). Bargy. Barodet. Barthou. Bartissol. Bastid (Adrien). Baulard. Beauquier. Bérard. Berger (Georges) (Seine). Bertrand. Bézine. Bizarelli. Bizot. Bizouard-Bert. Blanc (Pierre). Boissy-d'Anglas. Bonnefoy-Sibour. Bony-Cisternes. Borriglione. Boucher (Vosges). Boudenoot. Boudeville. Bouge. Boulanger-Bernet. Boullay. Bourgeois (Jura). Bourgeois (Léon) (Marne). Bouthier de Rochefort. Bovier-Lapierre. Boysset. Braud. Breton. Briens. Brisson (Henri). Brousse (Emile). Brugnot. Brunier. Bully. Burdeau. Buvignier.

Rectifications aux scrutins de la séance du 25 novembre 1889.

M. Michau (Nord), porté comme s'étant abstenu dans le scrutin sur l'urgence de la proposition de M. Maxime Lecomte, déclare avoir voté « pour ».

(a) Source numérisée

```
<!-- ANNEXES -->
<back>
   <head>Annexes au procès-verbal de la séance du <date when="1889-11-26">mardi 26 novembre 1889</date>.</head>

   <div xml:id="vot18891126">
      <!-- VOTE 1 -->
      <div xml:id="vot18891126_vot1" type="voting" corresp="#discussion7ebureau">
         <head>
            <label>SCRUTIN</label>
            <note><seg>Sur les conclusions du <num>7e</num> bureau tendant à l'annulation des opérations électorales de la
<placeName ref="#lieu_ID"><num>1re</num> circonscription de l'arrondissement de Lorient (Morbihan)</placeName>.</seg></note>
         </head>
         <!-- Détail du vote -->
         <desc>
            <measure type="nbvotants" quantity="506">Nombre des votants <num>506</num></measure>
            <measure type="maj" quantity="254">Majorité absolue <num>254</num></measure>
            <measure type="ayes" quantity="330">Pour l'adoption <num>330</num></measure>
            <measure type="noes" quantity="176">Contre <num>176</num></measure>
         </desc>

         <note type="result"><seg>La <orgName ref="#org_ID">Chambre des députés</orgName> a adopté.</seg></note>

         <floatingText><body><div><pb n="192"/></div></body></floatingText>

         <!-- Liste des votants -->
         <note type="voterslist">
            <desc>Ont voté pour :</desc>
            <seg> MM.<persName ref="#pers_ID">Abeille</persName>. <persName ref="#pers_ID">Arène (Emmanuel)</persName>. <persName
ref="#pers_ID">Armez</persName>. <persName ref="#pers_ID">Arribat</persName>. <persName ref="#pers_ID">Audifred</persName>. <persName ref
="#pers_ID">Aynard (Edouard)</persName>.</seg>
               <!-- [...] -->
         </note>
      </div>

      <!-- RECTIFICATIONS -->
      <div corresp="#vot18891125" type="rectification">
         <head>Rectifications aux scrutins de la séance du <date>25 novembre 1889</date>.</head>
         <note corresp="#vot18891125_vot1"><seg><persName ref="#pers_ID">M. Michau</persName> <placeName
ref="#lieu_ID">(Nord)</placeName>, porté comme s'étant abstenu dans le scrutin sur l'urgence de la proposition de <persName
ref="#pers_ID">M. Maxime Lecomte</persName>, déclare avoir voté « pour ».</seg></note>
         <!-- [...] -->
      </div>
   </div>
</back>
```

(b) Modèle d'encodage

Figure – Séance parlementaire du 26 novembre 1889 - votes, liste des votants, rectifications (extrait annexes)

STRUCTURE GÉNÉRALE DES FICHIERS XML TEI

```xml
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0" xml:id="FR_3R_5L" xml:lang="fr">

   <!-- Métadonnées du corpus (non développées) -->
   <teiHeader>
      <fileDesc>
         <titleStmt>
            <title></title>
         </titleStmt>
         <publicationStmt>
            <p></p>
         </publicationStmt>
         <sourceDesc>
            <p></p>
         </sourceDesc>
      </fileDesc>
   </teiHeader>

   <!-- Données liées et informations contextuelles -->
   <standOff>
      <listPerson>
         <person></person>
      </listPerson>

      <listOrg>
         <org></org>
      </listOrg>

      <listPlace>
         <place></place>
      </listPlace>
   </standOff>

   <!-- Stockage du composant correspondant à la séance du 26 novembre 1889 -->
   <xi:include xmlns:xi="http://www.w3.org/2001/XInclude" href="FR_3R_5L_1889-11-26.xml"/>


   <!-- Stockage des autres composants du corpus de façon identique -->
   <!-- ... -->

</teiCorpus>
```

(a) Structure générale d'un fichier corpus

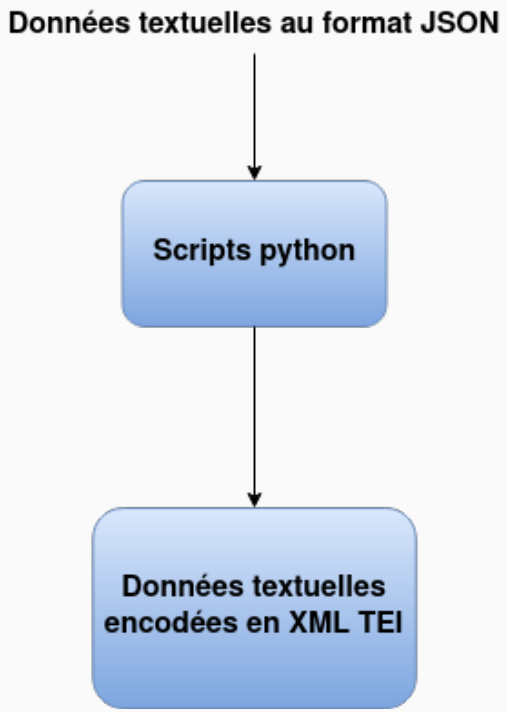
```xml
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="FR_3R_5L_1889-11-26" xml:lang="fr">

   <!-- Métadonnées du composant (non développées) -->
   <teiHeader>
      <fileDesc>
         <titleStmt>
            <title></title>
         </titleStmt>
         <publicationStmt>
            <p></p>
         </publicationStmt>
         <sourceDesc>
            <p></p>
         </sourceDesc>
      </fileDesc>
   </teiHeader>

   <text>

      <!-- Transcription du compte rendu -->
      <body>
         <div></div>
      </body>

      <!-- Annexes du compte rendu -->
      <back></back>

   </text>

</TEI>
```

(b) Structure générale d'un fichier composant

Figure – Encodages - Structure générale des fichiers corpus et composant

APPLIQUER L'ENCODAGE : AUTOMATISATION


```
19:
  activities:        []
  addresses:         []
  box:
    0:               220.12998972250773
    1:               282.1299897225077
    2:               561.8527187504491
    3:               80.4418437486525
  checked:           true
  comment:           "seg"
  id:                276
  ner_xml:           "<PER>M. Borie</PER> a <ACT>déjà fait partie des Assemblées\u2029législatives et satisfait aux conditions d</ACT>'<ACT>âge\u2029et de
                     nationalité</ACT> exigées par la loi.\u2029L'élection de la 2e circonscription de"
  origin:            "computer"
  parent:            261
  persons:
    0:               "M. Borie"
  text_ocr:          "M. Borie a déjà fait partie des Assemblées\nlégislatives et satisfait aux conditions d'âge\net de nationalité exigées par la loi."
  type:              "ENTRY"
```

DE JSON VERS XML-TEI

- Données textuelles dans les fichiers JSON = valeurs des clés ”text_ocr”
- Clés contenues dans des objets qui correspondent à la segmentation au niveau paragraphe
- Appliquer des règles de transformation sur les valeurs des clés ”text_ocr”

A LA RECHERCHE DE "FEATURES"

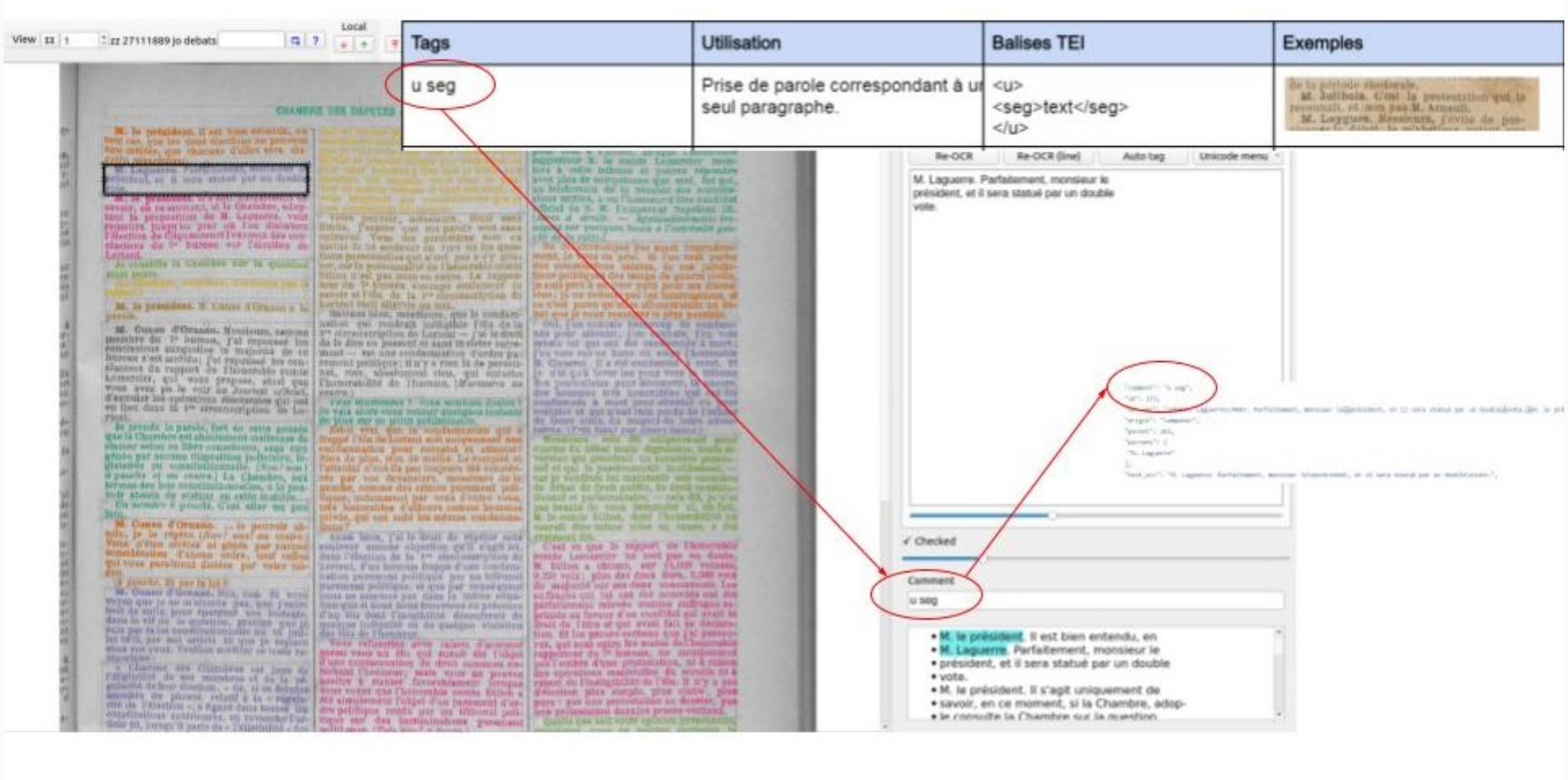


Figure – Du guide d'annotation à l'export JSON

RÉSULTATS


```
{
 "activities": [],
 "addresses": [],
 "box": [
  57.6116701150507,
  1710.0,
  560.3883298849495,
  50.07766597698992
 ],
 "checked": true,
 "comment": "u seg",
 "id": 305,
 "key": [
  0,
  1735
 ],
 "ner_xml": "<PER>M. Paul Déroulède</PER>. Je demande la pa-<0x2029>Tole",
 "origin": "computer",
 "parent": 269,
 "persons": [
  "M. Paul Déroulède"
 ],
 "text_ocr": "M. Paul Déroulede. Je demande la pa-\nrole.",
 "type": "ENTRY"
},
```


(a) Fichier JSON avec une prise de parole

```
def add_utterance(data):
    """
    Ajout de l'élément TEI "u" pour chaque boxe étiquetée "u" ou "u-beginning" et "u-end"
    :param data: dictionnaire contenant l'ensemble des données issues des JSON
    """
    for i in range(len(data)):
        if "comment" in data[i]:
            if re.search(r"\bu(?!-)\b", data[i]["comment"]):
                data[i]['text_ocr'] = "".join(['<u>', data[i]['text_ocr'], '</u>'])
            elif re.search(r"u-beginning", data[i]["comment"]):
                data[i]['text_ocr'] = "".join(['<u>', data[i]['text_ocr']])
            elif re.search(r"u-end", data[i]["comment"]):
                data[i]['text_ocr'] = "".join([data[i]['text_ocr'], '</u>'])
            else:
                pass
    return data
```

(b) Fonction add_utterance

```
<u><seg>M. Paul Déroulede. Je demande la parole.</seg></u>
```

(c) Résultat

Figure – Exemple d'application de la fonction add_utterance

CRÉATION DES FICHIERS COMPOSANTS

1. Récupération des données JSON stockées dans une variable ”data” sous forme d’un dictionnaire.
2. Gestion de l’encodage des changements de page.
3. Intégration des scripts contenant les règles de transformation dans la chaîne de traitement. Permettent d’encoder les données contenues dans la variable ”data”.
4. Création de l’élément <teiHeader> contenant les métadonnées :
    - Rédaction d’une partie du <teiHeader> à la main contenant les métadonnées ”fixes”
    - Construction de règles permettant de rechercher les métadonnées propres à chaque compte rendu, présentes dans la variable ”data”, par la suite intégrées au <teiHeader> préalablement établi
5. Création des fichiers XML valides.
6. Nettoyage des fichiers XML (gestion des césures).

Création du fichier corpus

1. Génération du fichier corpus.
2. Intégration des métadonnées.
3. Gestion des <xi :include>.

ANNOTATION SEMI-AUTOMATIQUE ET EXPORT EN XML-TEI

1. Annotation manuelle très consommatrice de temps et non exempte d'erreurs
   - Ajout de manière semi-automatique des étiquettes
   - Permettre aux utilisateurs de réutiliser les étiquettes et de créer leurs propres étiquettes (vocabulaire contrôlé ?)
2. Utiliser ces scripts pour proposer un export XML-TEI (outil SODUCO)
   - Ajouter les entités nommées (<persName>, <orgName>, etc.)

Demande de financement dans le cadre de Biblissima +

TOPIC MODELING ET WORD EMBEDDING

DÉCOMPOSITION EN VALEURES SINGULIÈRES

- On observe deux blocs de documents :
  - Mots positivements corrélés à la première composante : martin, paul, jules, léon, vicomte, seine, baron, comte...
  - Mots positivements corrélés à la seconde composante : monsieur, commission, ministre, projet, chambre, président, loi...

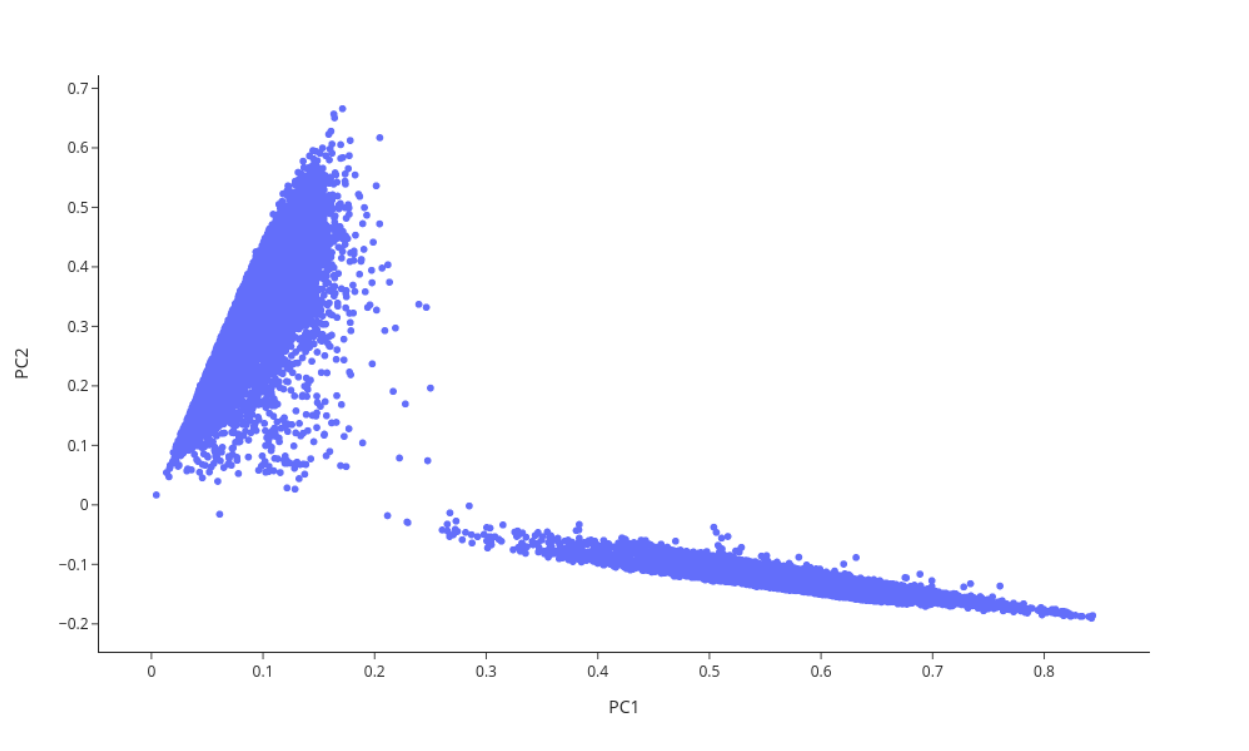


Figure – Décomposition en valeur singulière de la matrice de fréquence (+ tfidf)

ONT VOTÉ POUR :

MM. Abbatucci. Achard. Agniel. Allain-Targé. Allègre. Allemand. Amat. André (Jules). Andrieux. Anisson-Duperron. Anthoard. Arenberg (prince d'). Ariste (d'). Armez. Arnoult. Arrazat. Audiffred. Aulan (marquis d'). Azémar.

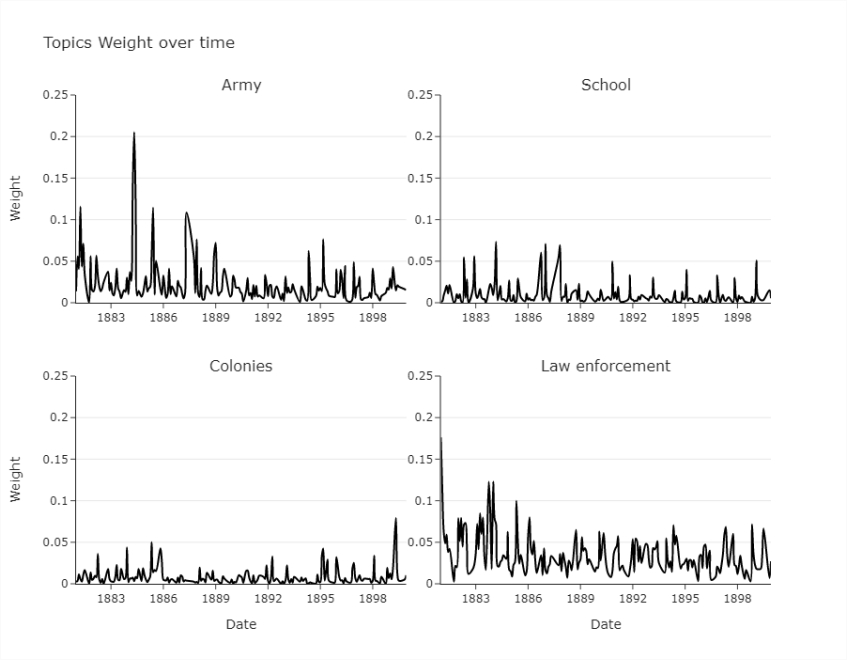
Baduel d'Oustrac. Baïhaut. Ballue. Bamberger. Bansard des Bois. Barascud. Barbedette. Bardoux. Barodet. Barthe (Marcel). Bastid (Adrien). Baudry-d'Asson (àe). Baury. Beauchamp (de). Beauquier. Beaussire. Bel (François). Belle. Bénazet. Benoist. Berger. Bergerot. Berlet. Bernard. Bernier. Bert (Paul). Bertholon. Bianchi. Bienvenu. Binachon. Bizarelli. Bizot de Fonteny. Blanc (Louis) (Seine). Blanc (Pierre) (Savoie). Blandin. Blin de Bourdon (vicomte). Bonnaud. Bonnet-Duverdier. Borriglione. Bosc. Bouchet. Boudeville. Boulard (Cher). Boulart (Landes). Bouquet. Bourgeois. Bousquet. Bouteille. Bouthier de Rochefort. Boyer (Ferdinand). Brame (Georges). Bravet. Breslay. Bresson. Brice (René). Brierre. Brisson (Henri). Brossard. Bruneau. Buyat.

Cadot (Louis). Caduc. Cantagrel. Carnot (Sadi). Casabianca (vicomte de). Casimir-Perier (Aube). Casimir-Perier (Paul) (Seine-Inférieure). Casse (Germain). Castagnède. Caurant. Caze. Cazeaux.

EXPLORER LES DÉBATS AVEC LA MODÉLISATION DE SUJETS

Topic 8	Topic 11	Topic 15
salaire	général	pari
question	commission	télégraphe
gouvernement	régiment	faire
jour	troupe	ingénieur
patron	monsieur	train
chambre	année	ligne
droit	jeune	chambre
syndicat	temps	personnel
délégué	faire	etat
monsieur	corps	administration
travail	soldat	employé
travaux	ministre	poste
ministre	homme	public
grève	loi	travaux
faire	an	service
mineur	guerre	agent
mine	service	ministre
loi	militaire	fer
compagnie	officier	chemin
ouvrier	armée	compagnie

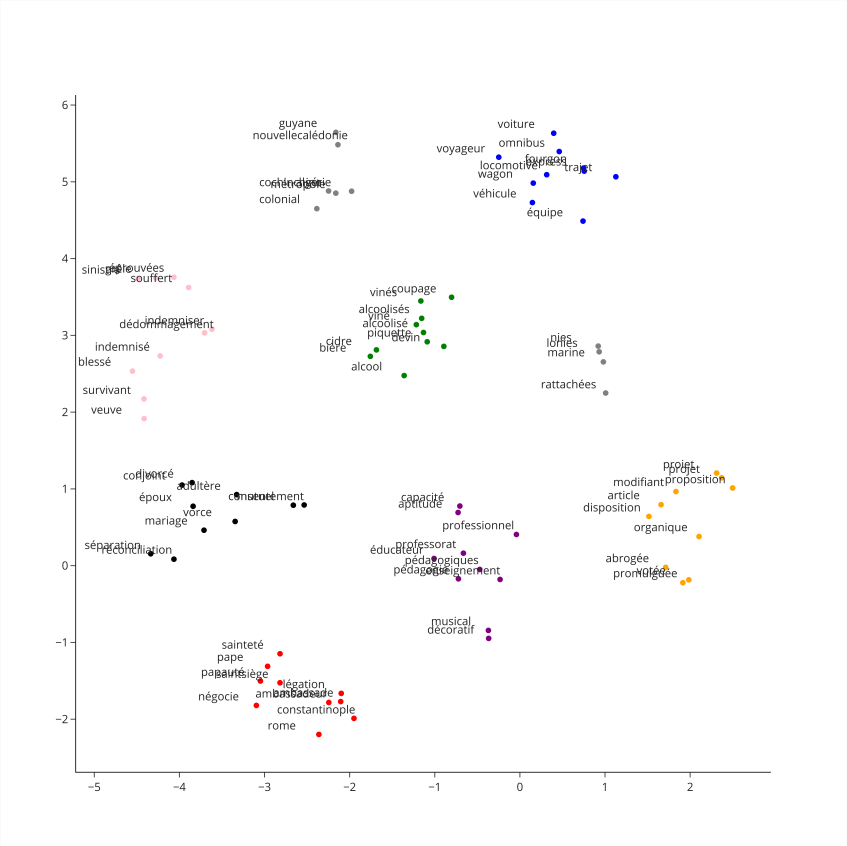
(a) 3 sujets parmi 40 : classe ouvrière (8), armée (11) et infrastructures (15)



(b) Evolution de quatre sujets au cours du temps

Figure – Résultats de la modélisation de sujets

LES PLONGEMENTS DE MOTS : WORD2VEC ET TOP2VEC



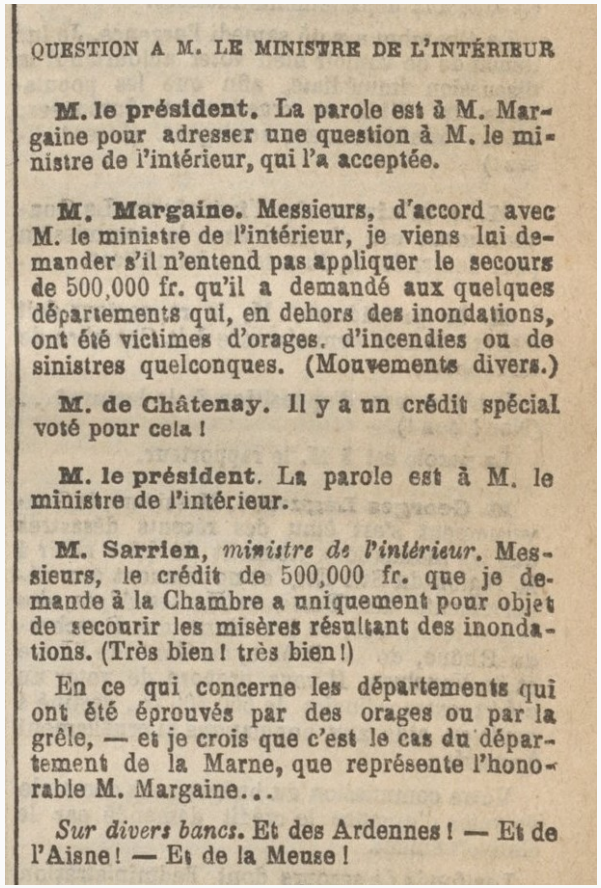
(a) Projection t-SNE des centroïdes des vecteurs (word2vec)

Cluster 55	Cluster 68	Cluster 70
victimes	divorce	enveloppes
inondations	epoux	timbres
secourir	mariage	poste
eprouvees	conjugal	postale
orages	divorces	timbre
sinistres	adultere	recepisses
grele	conjugale	postes
secours	remarier	postaux
venir	separation	telegraphes
infortunes	indissolubilite	colis
ravages	conjoints	fixe
miseres	mutuel	recouvrements
catastrophe	separations	graphes
evenements	mari	postales
repartition	mariages	taxe
incendies	femme	decide
soulager	conjoint	soit

(b) 3 clusters parmi les 113 obtenus avec top2vec : tempêtes (55), divorce (68) et poste (70)

Figure – Résultats obtenus avec word2vec et top2vec

Possibilité de calculer la cosine similarity entre un topic et des documents



QUESTION A M. LE MINISTRE DE L'INTÉRIEUR

M. le président. La parole est à M. Margaine pour adresser une question à M. le ministre de l'intérieur, qui l'a acceptée.

M. Margaine. Messieurs, d'accord avec M. le ministre de l'intérieur, je viens lui demander s'il n'entend pas appliquer le secours de 500,000 fr. qu'il a demandé aux quelques départements qui, en dehors des inondations, ont été victimes d'orages, d'incendies ou de sinistres quelconques. (Mouvements divers.)

M. de Châtenay. Il y a un crédit spécial voté pour cela !

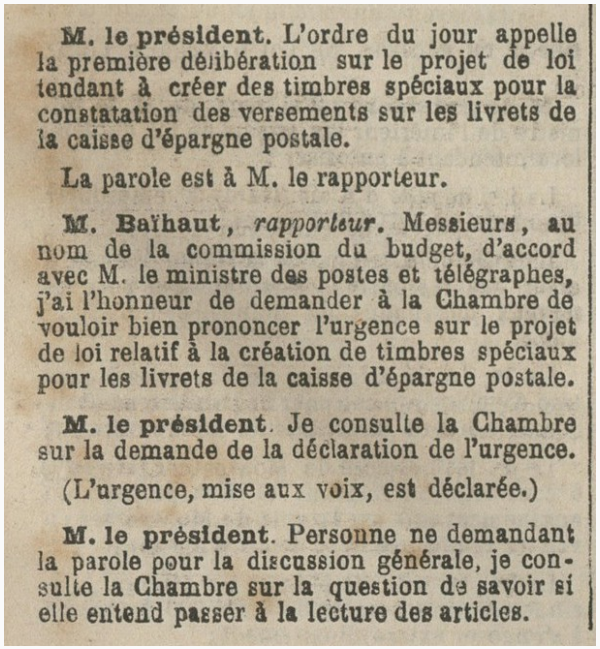
M. le président. La parole est à M. le ministre de l'intérieur.

M. Sarrien, *ministre de l'intérieur.* Messieurs, le crédit de 500,000 fr. que je demande à la Chambre a uniquement pour objet de secourir les misères résultant des inondations. (Très bien ! très bien !)

En ce qui concerne les départements qui ont été éprouvés par des orages ou par la grêle, — et je crois que c'est le cas du département de la Marne, que représente l'honorable M. Margaine...

Sur divers bancs. Et des Ardennes ! — Et de l'Aisne ! — Et de la Meuse !

(a) cos sim = 0.73



M. le président. L'ordre du jour appelle la première délibération sur le projet de loi tendant à créer des timbres spéciaux pour la constatation des versements sur les livrets de la caisse d'épargne postale.

La parole est à M. le rapporteur.

M. Baïhaut, *rapporteur.* Messieurs, au nom de la commission du budget, d'accord avec M. le ministre des postes et télégraphes, j'ai l'honneur de demander à la Chambre de vouloir bien prononcer l'urgence sur le projet de loi relatif à la création de timbres spéciaux pour les livrets de la caisse d'épargne postale.

M. le président. Je consulte la Chambre sur la demande de la déclaration de l'urgence.

(L'urgence, mise aux voix, est déclarée.)

M. le président. Personne ne demandant la parole pour la discussion générale, je consulte la Chambre sur la question de savoir si elle entend passer à la lecture des articles.

(b) cos sim = 0.81

Figure – Résultats obtenus avec word2vec et top2vec

POUR EN SAVOIR PLUS

- Nicolas Bourgeois, Aurélien Pellet, Marie Puren. ”Using Topic Generation Model to explore the French Parliamentary Debates during the early Third Republic (1881-1899)”. (hal-03526254v2)
- Marie Puren, Nicolas Bourgeois, Aurélien Pellet, Pierre Vernus, Fanny Lebreton. ”Between History and Natural Language Processing : Study, Enrichment and Online Publication of French Parliamentary Debates of the Early Third Republic (1881-1899)”. ParlaCLARIN III at LREC2022 - Workshop on Creating, Enriching and Using Parliamentary Corpora, Jun 2022, Marseille, France. (hal-03623351)

MERCI POUR VOTRE ATTENTION !

Aurélien Pellet : aurelien.pellet@epitech.eu

Marie Puren : marie.puren@epita.fr

Répertoire Github du projet : https://github.com/mpuren/agoda
Mémoire de Fanny Lebreton :
https://github.com/FannyLbr/Memoire-AGODA-TNAH2022