

Congrès AFDC

Intelligence artificielle et jurisprudence constitutionnelle : le traitement automatique des langues au service de la compréhension et de la prévision du contentieux électoral devant le Conseil constitutionnel

Présentation par Romain Rambaud (PR Droit, UGA) et Fang Zhao (post-doctorant Informatique, LIG)

Publications du projet JADE

1. R. Rambaud, A. Hafsaoui, C. Bligny, « Une justice algorithmique pour les élections politiques ? », *AJDA*, 2023, p. 1323.
2. R. Rambaud, C. Bligny, A. Hafsaoui, S. Cottin, « Justice algorithmique des élections (JADE) : une nouvelle base de données du contentieux des élections législatives devant le Conseil constitutionnel », *Jurimétrie*, 2025
- 2-1. C. Bligny, F. Letué, M.-J. Martinez, R. Rambaud, A. Hafsaoui, « Cross-Referenced Data on Electoral Disputes and French Legislative Election Results », *Journal of Open Humanities Data*, 2025, vol. 11, 34.

3 et 4. R. Rambaud, 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, RFDC, n°140, janv. 25,

Partie 2/2 – Sincérité du scrutin et critère de l'écart de voix, RFDC, n°141, mars 25

5. R. Rambaud, F. Letué, M-J. Martinez, F. Gougou, C. Bligny, « Le Conseil constitutionnel, juge électoral, est-il politisé ? », Colloque Toulon, octobre 2025, en cours de publication, Confluence des droits

6. (en cours / ANR) F. Zhao, R. Rambaud, A. Hafsaoui, « Projet JADE : le traitement automatique des langues au service de l'annotation semi-automatique des décisions du Conseil constitutionnel », Jurimétrie ?

Problématiques et hypothèses

- **Problématique** : quelle rationalité = notion de sincérité du scrutin, écart de voix ?
- **Hypothèse / contentieux des élections politiques** : contentieux ancien, standardisé, normé, chiffré, donc plus réceptif aux outils de la justice algorithmique....
- **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).

Résultats obtenus : la rationalité de la structuration générale du contentieux électoral (élections législatives)

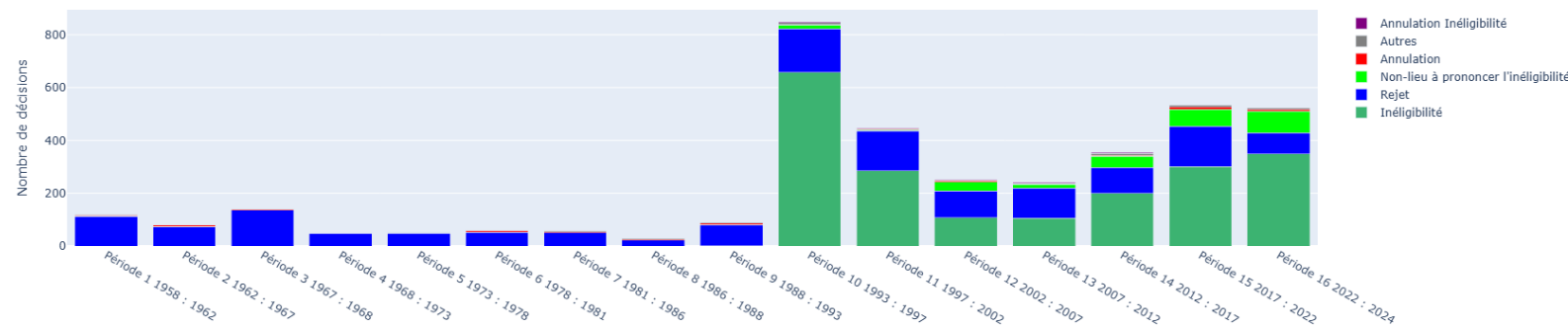
Effectifs AN par solution du Conseil Constitutionnel

Solution AN	Effectif
Inéligibilité	2003
Rejet	1479
Non-lieu à prononcer l'inéligibilité	254
Annulation	67
Non-lieu à statuer	41
Annulation Inéligibilité	16
Rectification erreur matérielle	7
Enquête	6
Irrecevabilité	3
Annulation non-lieu à statuer qpc	2
Approbation du compte	1
Inéligibilité Rejet qpc	1
Rejet non-lieu à statuer qpc	1

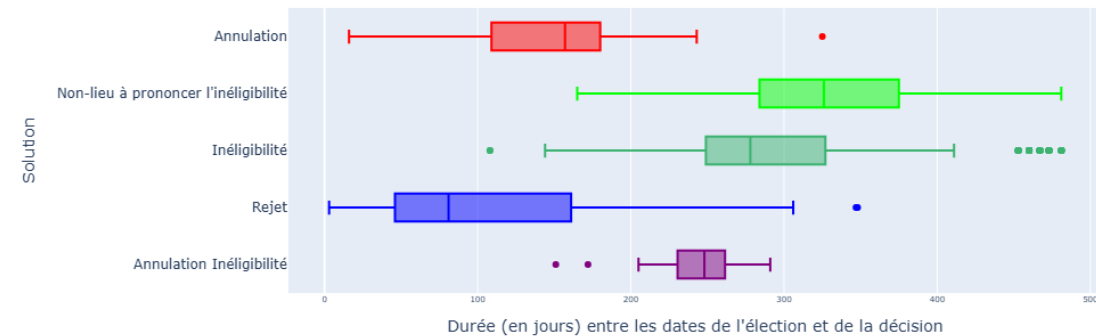
Effectifs AN après regroupement pour le projet jade

Solution AN Jade	Effectif
Inéligibilité	2004
Rejet	1483
Non-lieu à prononcer l'inéligibilité	255
Annulation	70
Autres	54
Annulation Inéligibilité	15

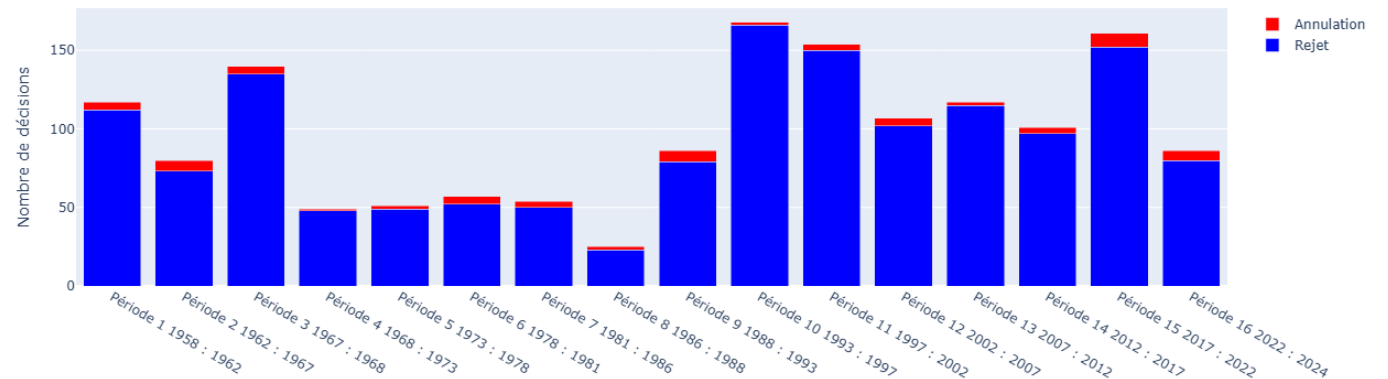
Nombre de décisions AN par période de législature et par solution



Répartition des durées par solution



Visualisation des solutions pour le contentieux des protestations électorales



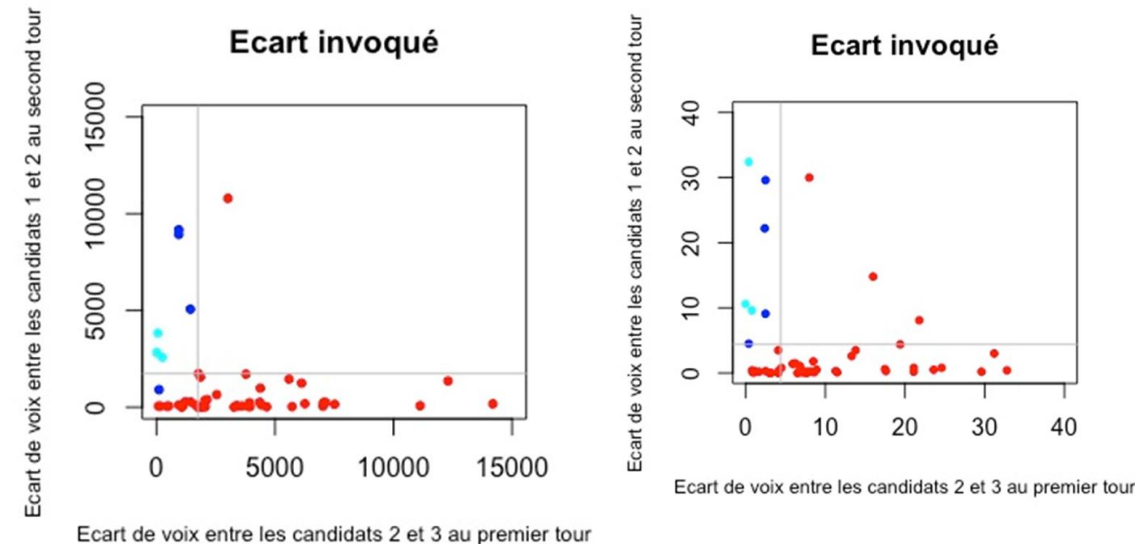
Résultats obtenus : la rationalité de la notion de sincérité du scrutin et de l'écart de voix (annulations)

- **Arbitraire ?** Robert (2000), Bernard Maligner (2013)
- **Immoral ?** Demichel (1973), Maligner (2017), Rousseau (2023)
- **Incohérent et trompeur ?** Franck (2010 et 2017)
- **Rationnel ?** Prunier (2003), Rambaud (2017 et 2020), JADE (2024-2025)

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

Annulation problèmes vote électronique : n° 2022-5813/5814 et n° 2022-5760 AN du 20 janvier 2023

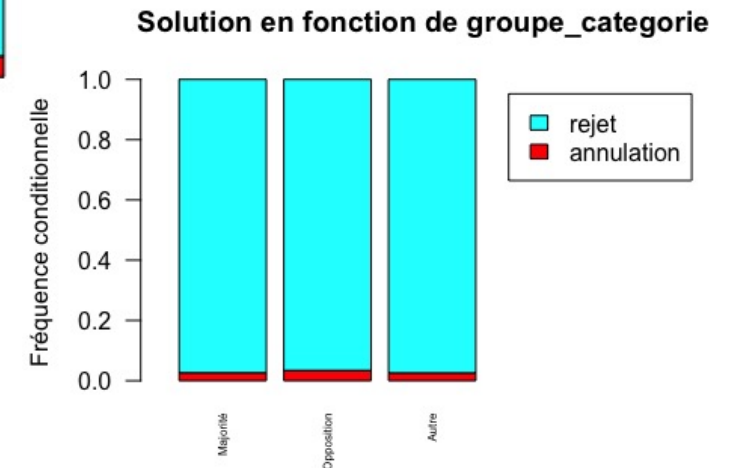
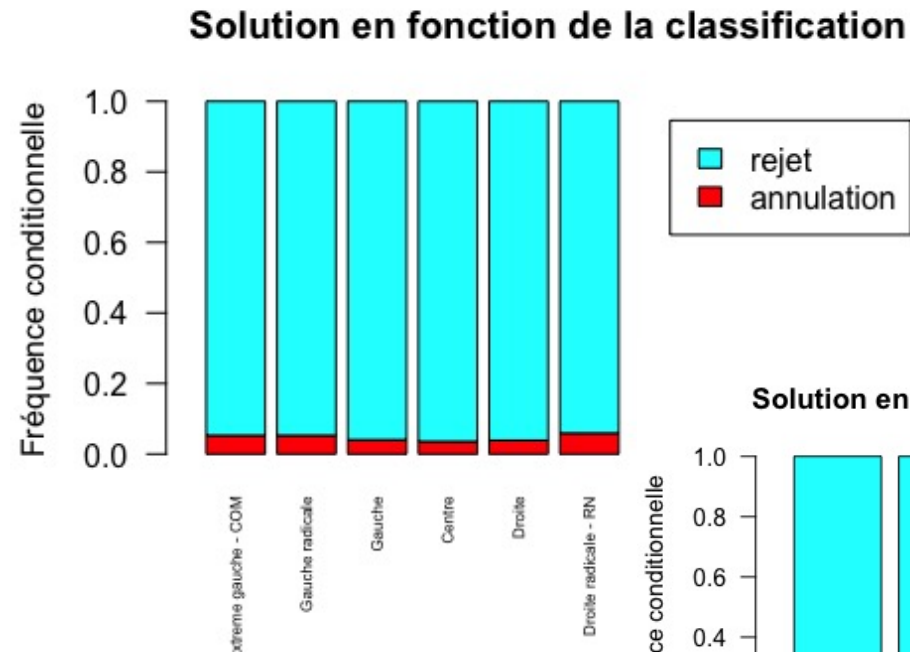
Inéligibilité candidat RN : n°2024-6341 du 13 février 2025



Résultats obtenus : l'absence de politisation manifeste du Conseil constitutionnel

- **Le Conseil constitutionnel politisé ?** Franck (2010) Fontaine (2023)

	annulation	rejet	Total
Extrême gauche - COM	5.3	94.7	100.0
Gauche radicale	5.3	94.7	100.0
Gauche	4.0	96.0	100.0
Ecologistes	0.0	100.0	100.0
Centre	3.6	96.4	100.0
Droite	3.9	96.1	100.0
Droite radicale - RN	5.9	94.1	100.0
Extrême Droite	0.0	100.0	100.0
Divers	0.0	100.0	100.0

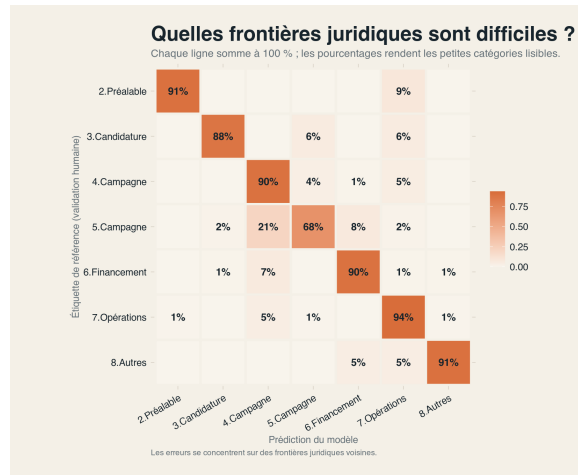


Le Conseil constitutionnel est-il « mainstream ? » (R. Franck) Entre 3 et 4 % en temps normal, entre 5 et 6 % pour les extrêmes ?

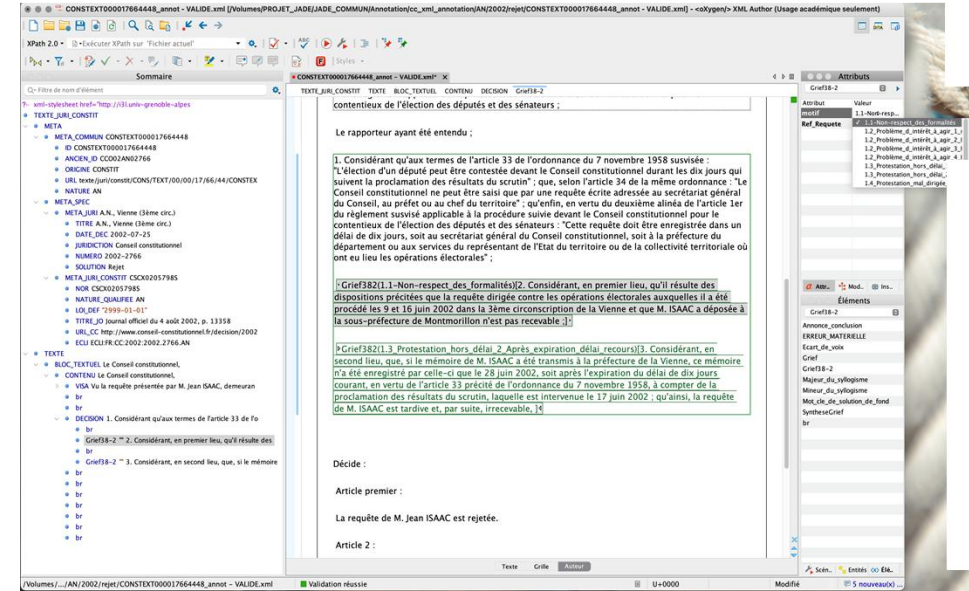
Taux d'annulation global : 4%, test de Fisher exact : 0.931

JADE et IA : les expérimentations du projet JADE en TAL

Objectif : lecture automatique des décisions du Conseil constitutionnel afin d'enrichir la base de données des griefs et réaliser des analyses de justice algorithmique



Catégorie	Précision	Rappel	Support moyen
2. Préalable	0.93	0.90	2
3. Candidature	0.89	0.87	3
4. Campagne	0.87	0.90	45
5. Campagne	0.81	0.68	13
6. Financement	0.92	0.90	20
7. Opérations	0.91	0.94	35
8. Autres	0.92	0.91	4



A	B	C	D	E	F
Vérification humaine	Remarque / état de l'annotation	Annotation alternative proposée	Fichier	Annotation	Texte
2	Oui	un considérant annoté sur 3	IAN1963:annotation.CONSTEXT00017665174_annot.xml	7 Opérations	1. Considérant que le procès-verbal des opérations
3	Oui		IAN1967:annotation.CONSTEXT00017665356_annot.xml	7 Opérations	4. Considérant qu'il n'est pas contesté que dans
4	Oui		IAN1967:annotation.CONSTEXT00017665356_annot.xml	7 Opérations	5. Considérant que, sans qu'il soit besoin d'avoir
5	Oui	considéran	IAN1968:annotation.CONSTEXT00017665568_annot.xml	7 Opérations	2. Considérant que, dans le double bureau de
6	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	3. Considérant qu'en pareil cas, il convient de re
7	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	4. Considérant qu'en soulignant un moyen tiré de
8	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	1. Considérant qu'aucune disposition de la loi n
9	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	4. Considérant que c'est également à tort que le
10	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	5. Considérant que c'est également à bon droi
11	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	7. Considérant que, de ce qui précède, il résulte
12	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	14. Considérant qu'aucune disposition de la loi n
13	Oui	considéran	IAN1978:annotation.CONSTEXT00017665825_annot.xml	7 Opérations	15. Considérant que la requérante n'apporte pas
14	Oui	considéran	IAN1981:annotation.CONSTEXT00017665942_annot.xml	7 Opérations	4. Considérant que les requérants font valoir que
15	Non	considéran	IAN1981:annotation.CONSTEXT00017665942_annot.xml	7 Opérations	1. Considérant qu'à la suite du scrutin du 21 juin
16	Oui	considéran	IAN1988:annotation.CONSTEXT00017667175_annot.xml	7 Opérations	2. Considérant, en premier lieu, qu'il résulte des
17	Non	considéran	IAN1988:annotation.CONSTEXT00017667175_annot.xml	4 Campagne	3. Considérant que ce dernier terme élève une r
18	Oui	considéran	IAN1998:annotation.CONSTEXT00017667910_annot.xml	4 Campagne	1. Considérant qu'aux termes du second alinéa c
19	Oui	considéran	IAN1998:annotation.CONSTEXT00017667910_annot.xml	4 Campagne	2. Considérant que la chaîne de télévision 'Cana
20	Oui	considéran	IAN1999:annotation.CONSTEXT00017667949_annot.xml	7 Opérations	2. Considérant que l'examen des feuilles de poi
21	Oui	considéran	IAN2002:annotation.CONSTEXT00017664509_annot.xml	6 Financement	1. Considérant qu'aux termes du deuxième alinéa
22	Oui	considéran	IAN2002:annotation.CONSTEXT00017664522_annot.xml	6 Financement	4. Considérant qu'il n'est établi ni que le coût de
23	Oui	considéran	IAN2002:annotation.CONSTEXT00017664522_annot.xml	6 Financement	5. Considérant qu'il n'est pas davantage établi q
24	Non	considéran	IAN2002:annotation.CONSTEXT00017664522_annot.xml	6 Financement	6. Considérant qu'il n'y a lieu, dès lors, ni de re
25	Oui	considéran	IAN2003:annotation.CONSTEXT00017664553_annot.xml	6 Financement	3. Considérant qu'en l'espèce l'objet de la réquisi
26	Oui	considéran	IAN2003:annotation.CONSTEXT00017664553_annot.xml	6 Financement	6. Considérant que des affiches signées 'PCF B
27	Oui	considéran	IAN2003:annotation.CONSTEXT00017664553_annot.xml	6 Financement	7. Considérant qu'il résulte de ce qui précède q
28	Oui	considéran	IAN2012:annotation.CONSTEXT00027090310_annot.xml	6 Financement	5. Considérant que, par une décision en date du
29	Non	considéran	IAN2012:annotation.CONSTEXT00027090310_annot.xml	7 Opérations	9. Considérant qu'il résulte de ce qui précède q
30	Oui	considéran	IAN2017:annotation.CONSTEXT000328390_annot.xml	4 Campagne	3. Aux termes de l'article L. 49 du code électora

L'annotation manuelle des griefs : un travail chronophage, difficile... et inutile si l'IA peut le faire à notre place ?

Classification de JADE

Classification inspirée des *tables analytiques du Conseil constitutionnel*

- **2. Préalables** : problèmes d'inscription sur les listes électorales, organisation des bureaux de vote, etc.
- **3. Candidatures** : inéligibilités, dépôt irrégulier, etc.
- **4. Campagne (moyens)** : affichage, réseaux sociaux, propagande tardive, réunions publiques, etc.
- **5. Campagne (pressions/manœuvres)** : pressions sur les électeurs, manœuvres frauduleuses, interventions irrégulières d'autorités, etc.
- **6. Financement** : recettes ou dépenses irrégulières, dons prohibés, dépassements de plafonds, etc.
- **7. Opérations électorales** : bulletins nuls, enveloppes irrégulières, dépouillement, procès-verbal, etc.
- **8. Autres** : cas non classables dans les catégories précédentes.

L'annotation manuelle des griefs dans Oxygen

Validation réussie

Premiers résultats d'annotation manuelle

Corpus : une centaine de décisions concernant des élections législatives de 2002 à 2022

Label	Objet	Occurrences
2.	Préalable	11
3.	Candidature	17
8.	Autres	22
5.	Campagne	66
6.	Financement	102
7.	Opérations	173
4.	Campagne	223
TOTAL		614

Table – répartition et labels des annotations suite à la manipulation de données

Face à la charge de l'annotation, comment le TAL peut-il aider ?

```
# 2. Load data according to the pre-fixed splits and prepare tokenization.
splits, rows_path, split_rows = load_split_rows(args.split_input)
labels, label_to_id = build_label_mapping(split_rows)

# Tokenization.
> tokenizer = AutoTokenizer.from_pretrained(...)
> _, dataloaders = build_dataloaders(...)

# 3. Initialize the classifier. The label set is inferred from the data.
> model = build_model(...)
# If freeze_encoder is True, only the classification head will be updated by the optimizer.
> if args.freeze_encoder:--

parameter_counts = count_parameters(model)
> loss_function, class_weights = build_loss_function(...)

device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
model.to(device)
loss_function.to(device)

# 4. Train on the fixed train split and monitor macro-F1 on dev after each epoch.
# Create output_dir if it doesn't exist to store temporary checkpoints
args.output_dir.mkdir(parents=True, exist_ok=True)
> training_result = train_model(...)

# 5. Evaluate the trained (and early-stopped) model on dev and test splits.
metrics_by_split = {}
predictions_by_split = {}
> for split_name in ("dev", "test"):--

# 6. Save the trained model, metrics, and per-instance prediction logs.
args.output_dir.mkdir(parents=True, exist_ok=True)
model_output_dir = args.output_dir / "model"
model_output_dir.mkdir(parents=True, exist_ok=True)

# Keep the training contract explicit in metrics.json so queue summaries and later audits
# do not need to reconstruct hyperparameters from shell history.
> write_json(...)
> write_json(...)
> for split_name, predictions in predictions_by_split.items():--

# The exported model directory is the final reusable checkpoint, unlike the temporary
# best_model.pt helper used only inside train_model().
save_checkpoint(model, tokenizer, model_output_dir)
```

1 Lecture des exemples validés

Le programme repart des cas déjà confirmés par l'équipe.

2 Apprentissage des régularités

Il repère les formulations qui reviennent dans les catégories juridiques.

3 Génération de propositions

Il produit ensuite des suggestions en série, à relire par un humain.

La machine prépare un brouillon ; elle ne décide pas à la place des juristes.

Notre système est-il une boîte noire ?

Systemes experts (Mots-clés)

Mécanisme : Règles explicites.

Bilan : 100% transparent, mais très rigide.

Notre système (Famille BERT)

Mécanisme : Modèle de langage spécialisé (ni règles, ni texte libre).

Bilan : Moins transparent, mais très cadré (erreurs quantifiables).

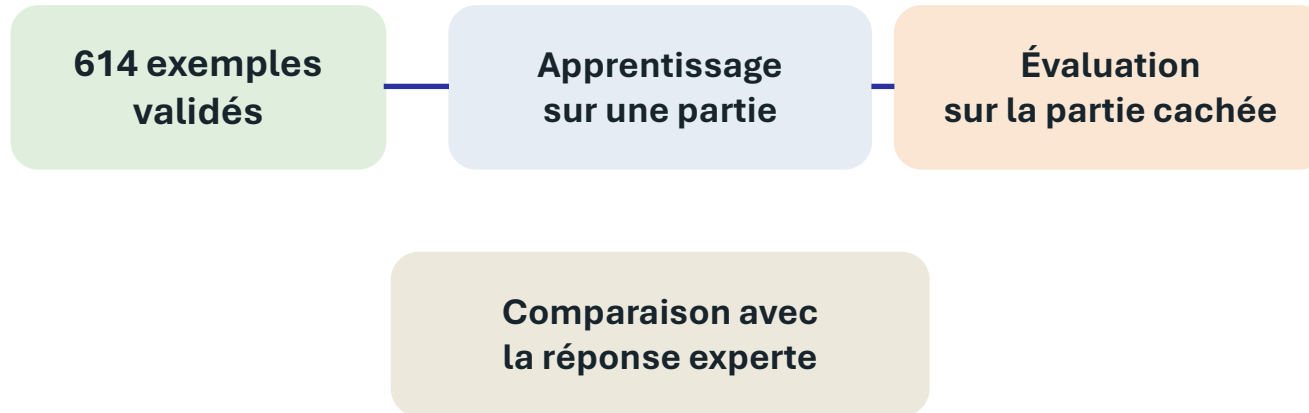
LLM Commerciaux (ex: ChatGPT)

Mécanisme : Modèle génératif fermé.

Bilan : Boîte noire totale, imprévisible.

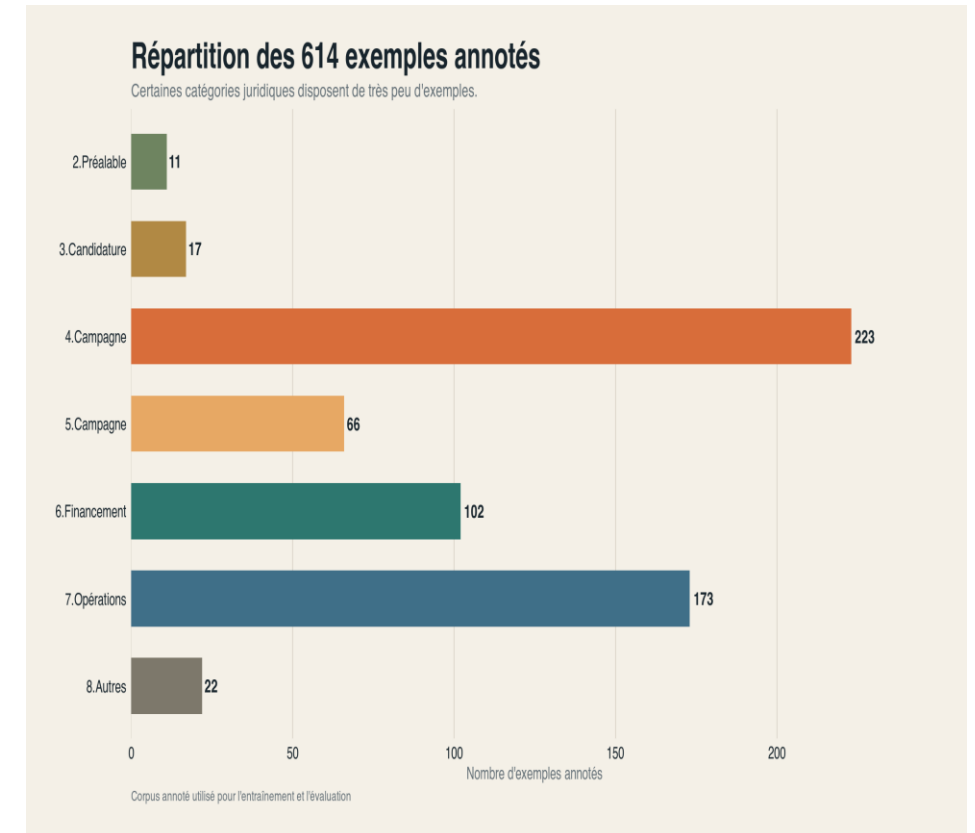
Comment évaluer équitablement les propositions de la machine ?

Principe de l'examen à livre fermé



Chaque exemple est testé quand la machine ne l'a pas vu.

Le support moyen indique combien d'exemples d'une catégorie apparaissent dans un test moyen.



Certaines catégories restent rares : les scores doivent toujours être lus avec le nombre d'exemples.

Quelles catégories sont fiables, lesquelles demandent une vigilance humaine ?

Catégorie	Précision	Rappel	Support moyen
2.Préalable	0.93	0.90	2
3.Candidature	0.89	0.87	3
4.Campagne (moyens)	0.87	0.90	45
5.Campagne (pressions/manœuvres)	0.81	0.68	13
6.Financement	0.92	0.90	20
7.Opérations	0.91	0.94	35
8.Autres	0.92	0.91	4

Précision Fiabilité

Quand la machine propose une catégorie, quelle est la probabilité que ce soit correct ?

Rappel Exhaustivité

Parmi les vrais cas d'une catégorie, combien la machine en retrouve-t-elle ?

Plus stable : 7.Opérations, 6.Financement, 4.Campagne (moyens).

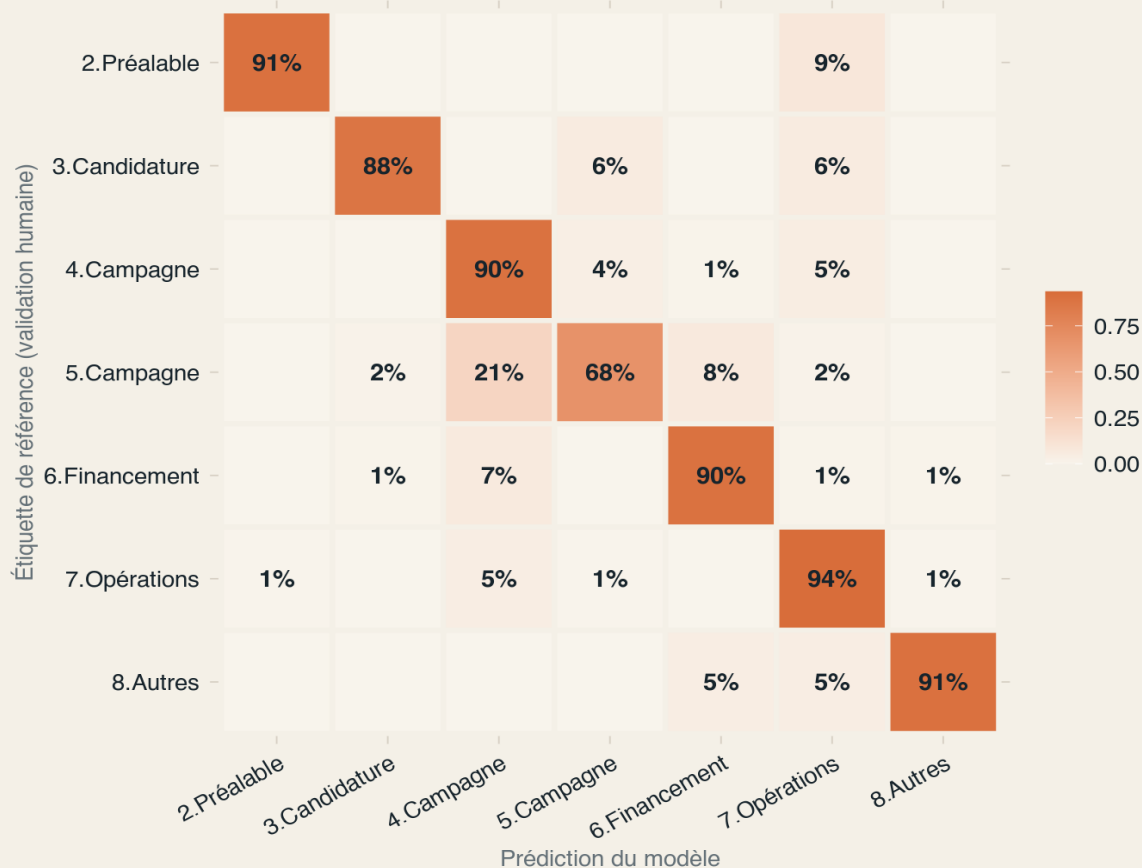
Point faible prioritaire : 5.Campagne (pressions/manœuvres).

Les petites catégories exigent plus de prudence.

Sur quelles frontières juridiques la machine hésite-t-elle ?

Quelles frontières juridiques sont difficiles ?

Chaque ligne somme à 100 % ; les pourcentages rendent les petites catégories lisibles.



Les erreurs se concentrent sur des frontières juridiques voisines.

5.Campagne (pressions/manœuvres) reste le cas le plus fragile

Environ 68 % des décisions de référence sont retrouvées comme 5.Campagne (pressions/manœuvres).

5.Campagne (pressions/manœuvres) glisse souvent vers 4.Campagne (moyens)

Environ 21 % des cas 5.Campagne (pressions/manœuvres) sont proposés comme 4.Campagne (moyens).

4.Campagne (moyens) et 7.Opérations se recourent

Ces erreurs signalent des frontières juridiques à relire, pas un échec aléatoire.

Point de vigilance : si l'outil propose 4 (moyens) ou 5.Campagne (pressions/manœuvres), la relecture humaine doit être renforcée.

Dans le workflow, comment les juristes valident-ils les propositions ?

1	Vérification humaine	Remarque / état de l'annotation	Annotation alternative proposée	Annotation	Texte
2	Oui	un considérant annoté sur 3		7 Opérations	1. Cons
3	Oui			7 Opérations	4. Cons
4	Oui			7 Opérations	5. Cons
5	Oui	considéran	considérants 3, 4, 5 ne sont pas annotés	7 Opérations	2. Cons
6	Oui	considérant 4 n'est pas annoté		7 Opérations	3. Cons
7	Oui	considéran	considéran	7 Opérations	1. Cons
8	Oui	considéran	considéran	7 Opérations	3. Cons
9	Oui	considéran	considéran	7 Opérations	4. Cons
10	Oui	considéran	considéran	7 Opérations	5. Cons
11	Oui	considéran	considéran	7 Opérations	7. Cons
12	Oui	considéran	considéran	7 Opérations	14. Cor
13	Oui	considéran	considéran	7 Opérations	15. Cor
14	Oui	considéran	considéran	7 Opérations	4. Cons
15	Non	considéran	ne devrait être annoté (Rappel des faits, ici rapp	7 Opérations	1. Cons
16	Oui	considéran	considéran	7 Opérations	2. Cons
17	Non	considérant 7 n'est pas annoté	3. Candidature : éligibilité / inéligibilité	4. Campagne	5. Cons
18	Oui			4. Campagne	1. Cons
19	Oui			4. Campagne	2. Cons
20	Oui	considérant 4 n'est pas annoté		7 Opérations	2. Cons
21	Oui	considéran	considéran	6 Financement	1. Con
22	Oui	considéran	considéran	6 Financement	4. Cons
23	Oui	considéran	considéran	6 Financement	5. Cons
24	Non	considéran	6. Financement ; ne devrait pas être annoté (Sy	4. Campagne	8. Cons
25	Oui	considéran	considéran	6 Financement	3. Cons
26	Oui	considéran	considéran	6 Financement	8. Cons
27	Non	considéran	6. Financement ; 4. Campagne (problème de p	4. Campagne	7. Cons
28	Oui	considéran	considéran	6 Financement	5. Cons
29	Non	considéran	ne devrait pas être annoté (Synthèse des griefs)	7 Opérations	9. Cons
30	Oui	considéran	considéran	4. Campagne	3. Aux

1 Proposition de la machine

Une catégorie ou une correction est préremplie pour accélérer la relecture.

2 Validation / correction humaine

Le juriste confirme, corrige ou refuse la proposition.

3 Donnée finale validée

La décision retenue repart dans la base comme donnée officielle.

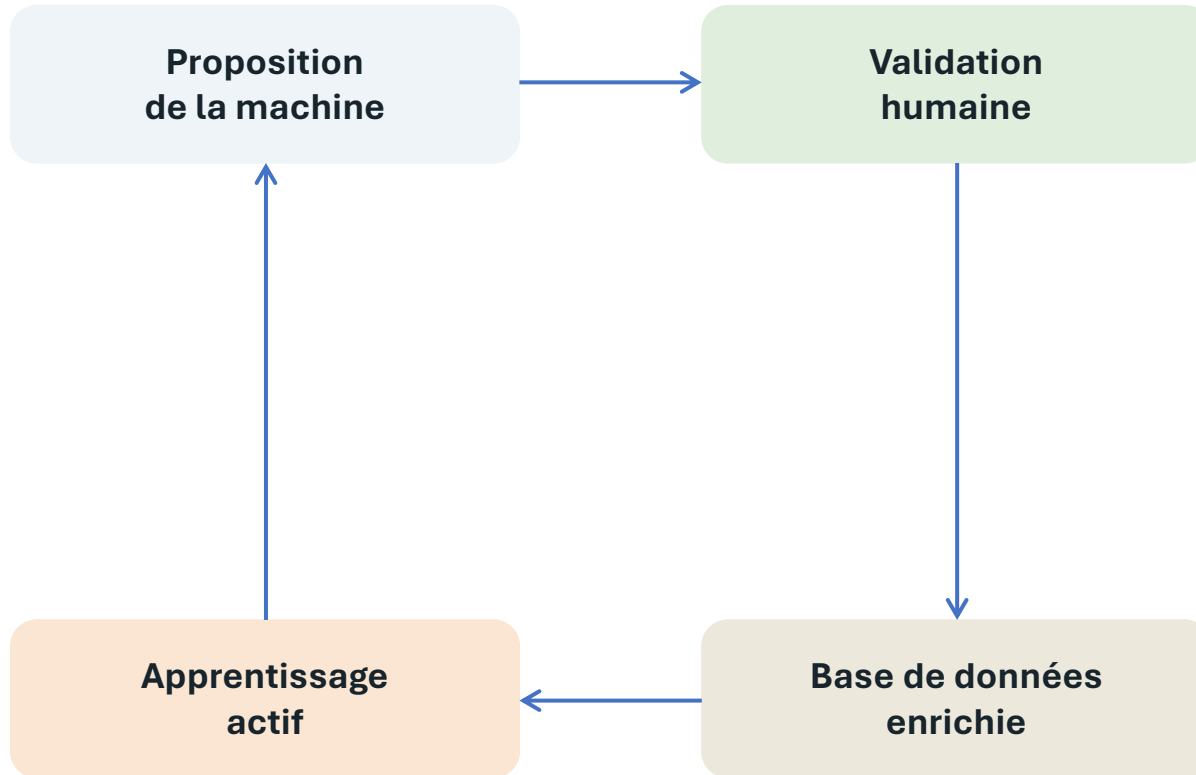
Machine : brouillon

Expert : décision

Base enrichie

Chaque correction alimente l'amélioration future : les juristes ne sont pas remplacés, ils entraînent l'outil à mieux respecter leur expertise.

En conclusion : une synergie pour le domaine juridique



Traçabilité Processus reproductible

Les résultats reposent sur une baseline vérifiable, pas sur une boîte noire incontrôlée.

Contrôle qualité Risques mesurés

Les zones d'erreur sont quantifiées et peuvent être priorisées pour la relecture.

Souveraineté humaine Décision finale

La qualification juridique reste toujours entre les mains des experts humains.

Au-delà de JADE, cette logique peut servir d'assistance pour d'autres tâches d'annotation et de contrôle de textes juridiques complexes.

Positionnement : outil d'assistance, pas système autonome de classification juridique.

Perspectives

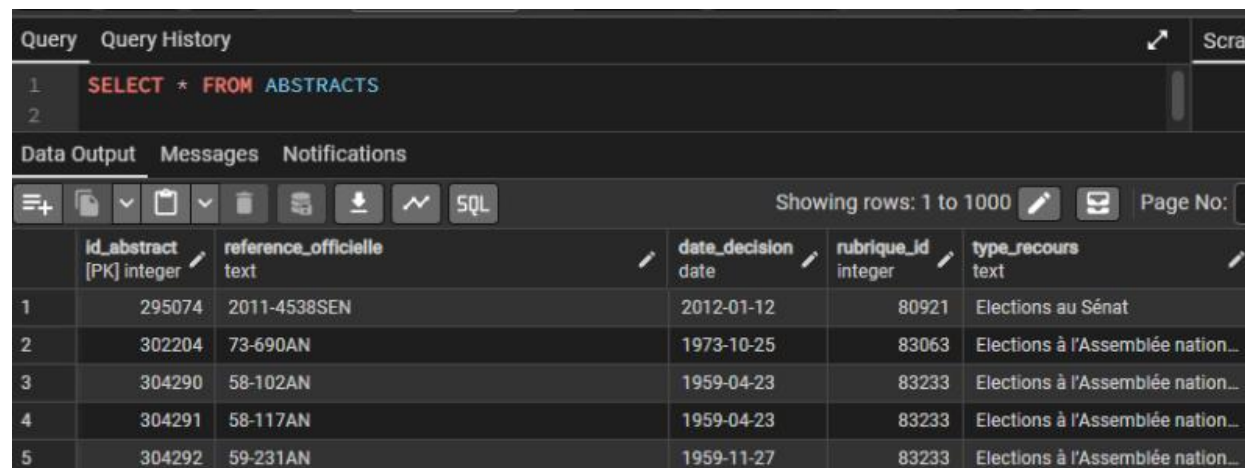
1 Descendre en sous-catégories dans la prédiction

La prédiction ne s'arrêtera pas aux 8 grandes catégories.

```
xml
<DECISION>
  <Grief objet="7.Opérations_électorales_11_Bulletins"
    qualification="1.Grief_NC_2_Rejet_au_fond_2_non_fondé">
    1. Il ressort des pièces du dossier que des bulletins au nom de Mme
    CARRIE,
      candidate soutenue par le Rassemblement national dans la deuxième
      circonscription
      de l'Ariège, ont été mêlés à ceux au nom de M. GARNIER, candidat
      également soutenu
      par ce parti, dans les bureaux de vote de la commune de Tarascon-sur-
      Ariège.
      136 bulletins au nom de Mme CARRIE ont été retrouvés dans l'urne et
      comptabilisés
      à bon droit comme nuls par la commission de recensement à l'issue du
      scrutin du
      12 juin 2022.
  </Grief>
</DECISION>
```

2 Exploitation des Abstracts

Structuration et cartographie avec les décisions.



The screenshot shows a database query interface. At the top, there's a 'Query' tab with a 'Query History' section. The query entered is 'SELECT * FROM ABSTRACTS'. Below the query, there's a 'Data Output' section with a table of results. The table has columns for 'id_abstract', 'reference_officielle', 'date_decision', 'rubrique_Jd', and 'type_recours'. The results show 5 rows of data.

	id_abstract [PK] integer	reference_officielle text	date_decision date	rubrique_Jd integer	type_recours text
1	295074	2011-4538SEN	2012-01-12	80921	Elections au Sénat
2	302204	73-690AN	1973-10-25	83063	Elections à l'Assemblée nation...
3	304290	58-102AN	1959-04-23	83233	Elections à l'Assemblée nation...
4	304291	58-117AN	1959-04-23	83233	Elections à l'Assemblée nation...
5	304292	59-231AN	1959-11-27	83233	Elections à l'Assemblée nation...

3 Raisonnement sur les CR joints

Tâches de raisonnement juridique sur les comptes rendus.



Merci !
Questions ?

Intelligence artificielle et jurisprudence constitutionnelle : le traitement automatique des langues au service de la compréhension et de la prévision du contentieux électoral devant le Conseil constitutionnel

Présentation par Romain Rambaud (PR Droit, UGA) et Fang Zhao (post-doctorant Informatique, LIG)