

Justice algorithmique des élections

Université Grenoble-Alpes, IDEX / IRGA UGA, 2022-2025

IERDJ, 2023-2026

Pré-proposition AAPG ANR 2023-2024 – Consortium UGA /
Université Nice Côte d'Azur – CERDACFF et GREDEG

22 novembre 2023

Romain Rambaud, Professeur de droit public, Université Grenoble-Alpes, CRJ

Caroline Bligny, Ingénieure de recherche CNRS, Université Grenoble-Alpes, LJK

Alya Hafsaoui, Doctorante contractuelle, Université Grenoble-Alpes, CRJ (distanciel)

Hypothèses du projet JADE

- **Constat** : montée en puissance de la « justice prédictive », de la « jurimétrie » mais peu de travaux en droit public (ex. Supra Legem)
- **Contentieux des élections politiques** : contentieux ancien, standardisé, normé, chiffré, donc plus réceptif aux outils de la justice algorithmique... (v. L. Godefroy, GREDEG). **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 (v. F. Marty, GREDEG)

Etape 1 : La constitution des bases de données (Data Paper en cours de réalisation)

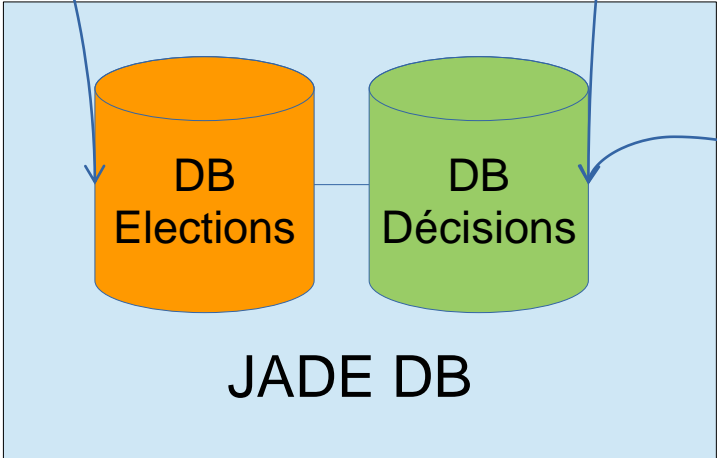
- Les bases de données Décisions : la problématique de la (non)préparation des données / métadonnées par les juridictions (XML – Zip Datagouv)
 - **Conseil constitutionnel** : Les balises sémantiques suivantes permettent de trier informatiquement les décisions selon : <ECLI> (numéro unique de la décision), <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** :

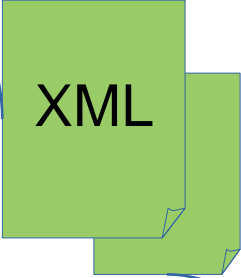
Code PJCA	Texte Code PCJA	Total décisions	Balises solution remplies			Balises solution non remplies	
			Nombre	%	Nb Métadonnées	Nombre	%
28-005	Dispositions générales	240	98	40,8	25	142	59,2
28-01	Présidentielle	13	7	53,8	5	6	46,2
28-02	Législatives	42	23	54,8	7	19	45,2
28-023	Européennes	37	10	27,0	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
Total	Elections politiques	3661	1118	30,5	-	2543	69,5

Construction

Résultats des élections,
data.gouv.fr

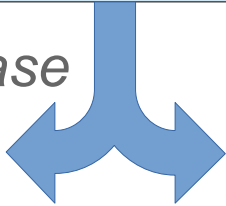


Décisions du conseil constitutionnel
data.gouv.fr

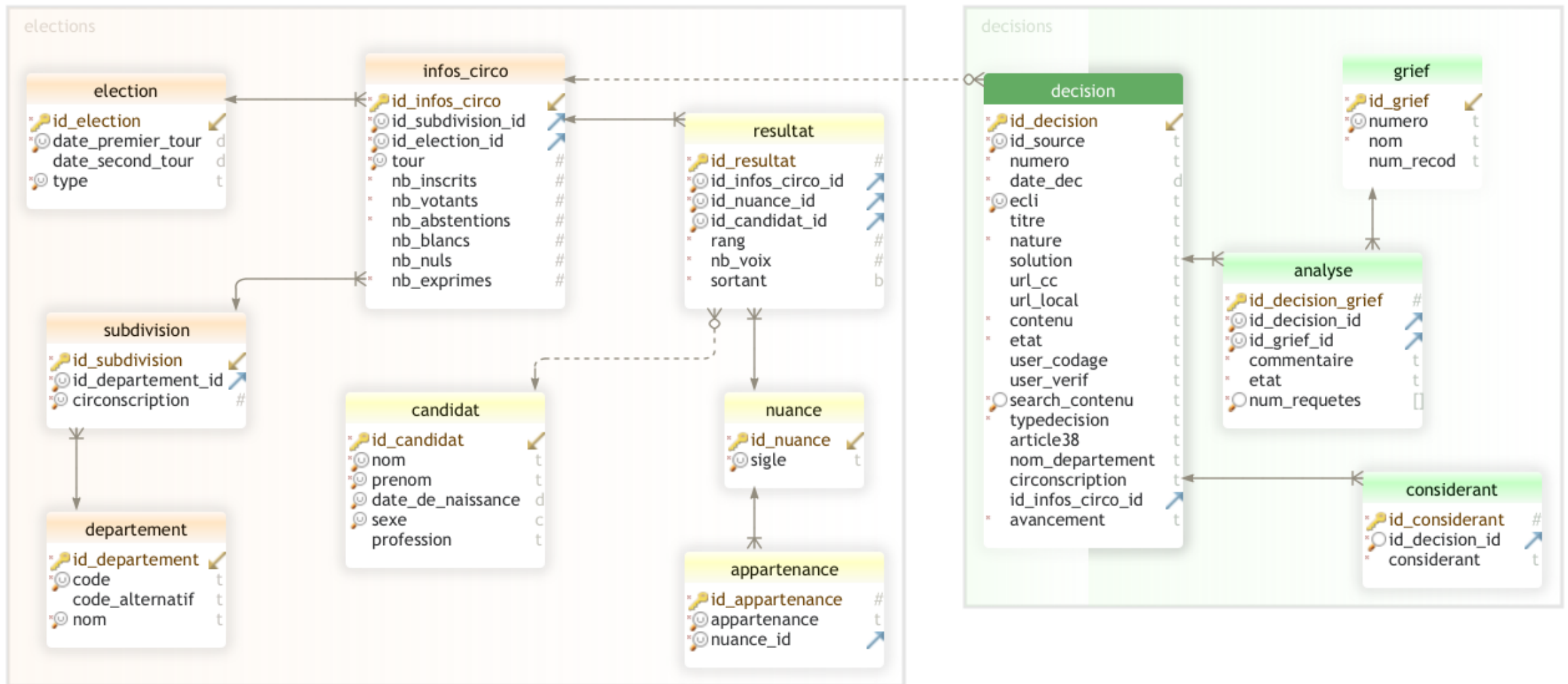


Annotation des décisions

Affichage, visualisation d'une partie de la base



Export de données consolidées à des fins d'analyse



Base de données JADE, Modèle Physique des Données

Quelques chiffres

- Au 04 mai dernier :
- 3888 décisions
- 3365 connectées à une élection

72	Annulation
21	Annulation - Inéligibilité
1	Approbation
1	Conformité et rejet
6	Enquête
1880	Inéligibilité
5	Irrecevabilité
223	Non lieu à prononcer l'inéligibilité
47	Non lieu à statuer
2	Non lieu à statuer sur la QPC et annulation
1	Null
8	Rectification d'erreur matérielle
1618	Rejet
1	Rejet - Conformité [QPC]
1	Rejet - non lieu à statuer [QPC]
1	Réformation

Visualisations des bases de données (décisions)

numero	date_dec	titre	solution	typedec...	url_cc
97-2434	1998-02-20	A.N., Haute-Vienne (3ème circ.)	Inéligibilité	unique	http://www.conseil-cor
97-2435	1998-02-20	A.N., Rhône (12ème circ.)	Inéligibilité	unique	http://www.conseil-cor
97-2436	1998-02-20	A.N., Rhône (12ème circ.)	Inéligibilité	unique	http://www.conseil-cor
97-2441	1998-02-20	A.N., Val-d'Oise (7ème circ.)	Inéligibilité	unique	http://www.conseil-cor
97-2443	1998-02-20	A.N., Val-d'Oise (6ème circ.)	Inéligibilité	unique	http://www.conseil-cor
97-2444	1998-02-20	A.N., Val-d'Oise (8ème circ.)	Inéligibilité	unique	http://www.conseil-cor
2012-4566	2012-07-13	A.N., Bouches-du-Rhône (3ème circ.)	Rejet	unique	http://www.conseil-cor
2012-4598	2012-12-07	A.N., Bouches-du-Rhône (3ème circ.)	Rejet	unique	http://www.conseil-cor
2012-4616	2012-11-29	A.N., Seine-Saint-Denis (8ème circ.)	Rejet	unique	http://www.conseil-cor
2012-4628	2012-12-14	A.N., Bouches-du-Rhône (4ème circ.)	Rejet	unique	http://www.conseil-cor
2012-4648	2012-07-20	A.N., Creuse (1ère circ.)	Rejet	unique	http://www.conseil-cor
97-2445	1998-02-20	A.N., Val-d'Oise (8ème circ.)	Inéligibilité	unique	http://www.conseil-cor
2012-4564	2012-07-20	A.N., Vaucluse (1ère circ.)	Rejet	unique	http://www.conseil-cor
2012-4569/4570/4571	2012-07-20	A.N., Vaucluse (3ème circ.)	Rejet	jonction	http://www.conseil-cor
2012-4629	2012-07-20	A.N., Val-d'Oise (8ème circ.)	Rejet	unique	http://www.conseil-cor
2012-4587	2012-11-20	A.N., Yvelines (11ème circ.)	Rejet	unique	http://www.conseil-cor
97-2450	1998-02-20	A.N., Paris (19ème circ.)	Inéligibilité	unique	http://www.conseil-cor
97-2451	1998-02-20	A.N., Paris (19ème circ.)	Inéligibilité	unique	http://www.conseil-cor
97-2499	1998-02-20	A.N., Seine-Saint-Denis (13ème circ.)	Inéligibilité	unique	http://www.conseil-cor

Visualisations des bases de données (candidats)

numero text	date_dec date	nom_departement text	circonscription text	nom text	prenom text	rang integer	nb_voix integer	contenu text
97-2504	1998-02-20	[ain]	[1ère circ.]	BOYON	JACQUES	2	14048	Le Conseil constitutionnel,
2002-2991	2003-02-06	[rhone]	[7ème circ.]	MILLON	CHARLES	2	7270	Le Conseil constitutionnel,
2002-2989	2003-02-06	[rhone]	[7ème circ.]	MILLON	CHARLES	2	7270	Le Conseil constitutionnel,
2002-2715	2002-07-25	[ain]	[4ème circ.]	VOISIN	MICHEL	1	28045	Le Conseil constitutionnel,
2002-2939	2003-02-27	[aisne]	[1ère circ.]	DOSIERE	RENE	1	14231	Le Conseil constitutionnel,
88-1059	1988-11-25	[aisne]	[5ème circ.]	JOURDAIN	DOMINIQUE	2	16029	Le Conseil constitutionnel,
88-1059	1988-11-25	[aisne]	[5ème circ.]	ROSSI	ANDRE	1	21282	Le Conseil constitutionnel,
88-1059	1988-11-25	[aisne]	[5ème circ.]	ROUSSEAU	MARCEL	3	5750	Le Conseil constitutionnel,
2002-2687/2741	2002-12-19	[allier]	[1ère circ.]	COLCOMBET	FRANCOIS	2	13532	Le Conseil constitutionnel,
2002-2793/2794/2795	2002-10-31	[allier]	[1ère circ.]	COLCOMBET	FRANCOIS	2	13532	Le Conseil constitutionnel,
2002-2791/2792	2002-10-31	[allier]	[1ère circ.]	COLCOMBET	FRANCOIS	2	13532	Le Conseil constitutionnel,
2002-2778	2002-11-07	[alpes-de-haute-provence]	[1ère circ.]	GALIZI	FRANCIS	2	10603	Le Conseil constitutionnel,
2002-2649/2663/2664	2002-11-07	[alpes-de-haute-provence]	[1ère circ.]	GALIZI	FRANCIS	2	10603	Le Conseil constitutionnel,
97-2175	1997-10-28	[hautes-alpes]	[2ème circ.]	OLLIER	PATRICK	1	8213	Le Conseil constitutionnel,
2012-4547	2012-07-13	[hauts-de-seine]	[7ème circ.]	OLLIER	PATRICK	1	23944	LE CONSEIL CONSTITUTIONNEL,

Visualisations des bases de données (écarts voix)

election_date	departement_text	circonscription_integer	voix_candidat_1_integer	voix_candidat_2_integer	ecart_de_voix_integer	decision_text	solution_text	url_cc_text
1958-11-23	hautes-alpes	1	4681	4406	275	58-43	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/5843an.fr
1958-11-23	ardeche	3	9536	9348	188	58-113	Annulation	http://www.conseil-constitutionnel.fr/decision/1959/58113an.l
1958-11-23	ardeche	3	9536	9348	188	59-206	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/59206an.l
1958-11-23	ardennes	2	10540	10316	224	58-28	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/5828an.fr
1958-11-23	aveyron	3	11682	11633	49	58-74	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/5874an.fr
1958-11-23	charente-maritime	5	11077	10917	160	58-96	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/5896an.fr
1958-11-23	nievre	2	11659	11199	460	58-9	Rejet	http://www.conseil-constitutionnel.fr/decision/1958/589an.htr
1958-11-23	bas-rhin	1	9947	9483	464	58-89	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/5889an.fr
1958-11-23	bas-rhin	1	9947	9483	464	58-47/72/122	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/5847_72_
1958-11-23	seine	7	8964	8615	349	61-234	Rejet	http://www.conseil-constitutionnel.fr/decision/1961/61234an.l
1958-11-23	seine-maritime	4	9312	9123	189	58-73/120	Rejet	http://www.conseil-constitutionnel.fr/decision/1959/5873_120
1962-11-18	aube	3	11429	10997	432	62-294	Rejet	http://www.conseil-constitutionnel.fr/decision/1963/62294an.l
1962-11-18	bouches-du-rhone	1	9519	9292	227	62-317	Rejet	http://www.conseil-constitutionnel.fr/decision/1963/62317an.l
1962-11-18	seine-maritime	4	10782	10540	242	62-265	Rejet	http://www.conseil-constitutionnel.fr/decision/1963/62265an.l
1967-03-05	haute-garonne	2	13803	13455	348	67-429	Rejet	http://www.conseil-constitutionnel.fr/decision/1967/67429an.l

Export pour les analyses statistiques

Caractéristiques décision
(solution, numéro, ..) + bilan
élection

v_bilan	
id_infos_circo_tf	#
id_infos_circo_t1	#
annee_election	#
date_premier_tour	d
date_second_tour	d
type_election	t
nom_dep	t
circonscription	#
tour_final	#
tf_nb_exprimees	#
tf_nb_votant	#
tf_nb_inscrit	#
tf_ecart1	#
tf_r1_nuance	t
tf_r1_nb_voix	#
tf_r1_sortant	b
tf_r1_civilite	c
tf_r2_nuance	t
tf_r2_nb_voix	#
tf_r2_sortant	b
tf_r2_civilite	c
t1_nb_exprimes	#
t1_nb_votant	#
t1_nb_inscrit	#
t1_ecart1	#
t1_ecart2	#
t1_r1_nuance	t
t1_r1_nb_voix	#
t1_r1_sortant	b
t1_r1_civilite	c
t1_r2_nuance	t
t1_r2_nb_voix	#
t1_r2_sortant	b
t1_r2_civilite	c
t1_r3_nuance	t
t1_r3_nb_voix	#
t1_r3_sortant	b
t1_r3_civilite	c

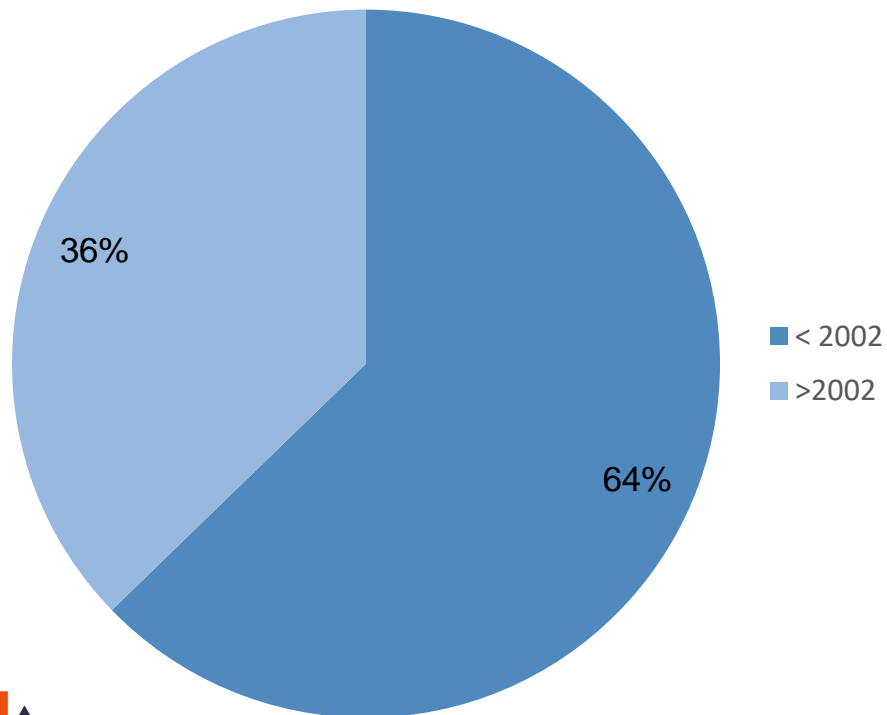
v_resultat	
id_resultat	#
id_infos_circo	#
nuance	t
nom_candidat	t
prenom_candidat	t
civilite_candidat	c
rang	#
rang_calc	#
nb_voix	#
sortant	b
id_candidat	#
id_nuance	#

v_infos_circo	
id_infos_circo	#
nom_dep	t
code_dep	t
tour	#
tour_inverse	#
circonscription	#
date_premier_tour	d
date_second_tour	d
type_election	t
nb_inscrits	#
nb_votants	#
nb_abstentions	#
nb_blancs	#
nb_nuls	#
nb_exprimes	#
id_subdivision	#
id_election	#

Etape 2 : Enrichissement des données - TAL en *machine learning* : les décisions de l'article 38 al. 2

- 1613 décisions de rejet au total

Répartition des Décisions



- Total des décisions labellisées après 2002(584)

Décisions de l'ordre de l'article 38 alinéa 2	Nature de la décision	
	Assemblée Nationale	Sénatoriale
Oui	327	35
Non	203	19

PRE-TRAITEMENT DES TEXTES

```
class TextDecisionProcessor(BaseEstimator, TransformerMixin):
#PROCESSUS DE NETOYAGE DU TEXTE POUR LA PERFORMANCE DU MODELE
    def __init__(self, n_jobs = -1,
                 remove_saut_ligne = True, #enlever les sauts de lignes
                 remove_decide = True, #enlever la partie décision
                 remove_special_char = True, #enlever les caractères spéciaux
                 remove_digits = True, # enlever les chiffres
                 remove_punctuation = True, #enlever la ponctuation
                 remove_stopword = True, #les mots inutiles
                 remove_people = True, #enlever les noms de personnes
                 remove_place = True, #les noms de lieux
                 text_lematization = True, # Lemmatisation du texte : peut -> pouvoir
                 remove_word = True, #enlever les mots < 3 caractères

                 #essayer d'enlever les mots qui permettent facilement de décider si décision est article 38 ou non
                 stopwords = stopwords.words('french') +
                 ["département", "L.O.-", "l.o.", "saisi", "saisir", "conseil", "constitutionnel", "alinéa", "article 38",
                 "alinéa 2", "deuxième", "mémoire", "mémoire", "défense", "réplique", "observations", "observation",
                 "rejeter", "irrecevable", "requête", "décision", "manifestement", "influence", "élection",
                 "résultat", "motiver", "grief", "enregistrer", "général", "voir", "circonscription", "circ", "date", "de",
                 "date", "secrétariat", "décision", "voir", "ordonnance", "contradictoire", "article", "de", "ensuit"

                 , "date", "secrétariat", "décision", "voir", "ordonnance", "contradictoire", "article", "de", "ensuit"

                 ):
                 punctuation = set(string.punctuation)
                 chiffres = set(list("0123456789")) :

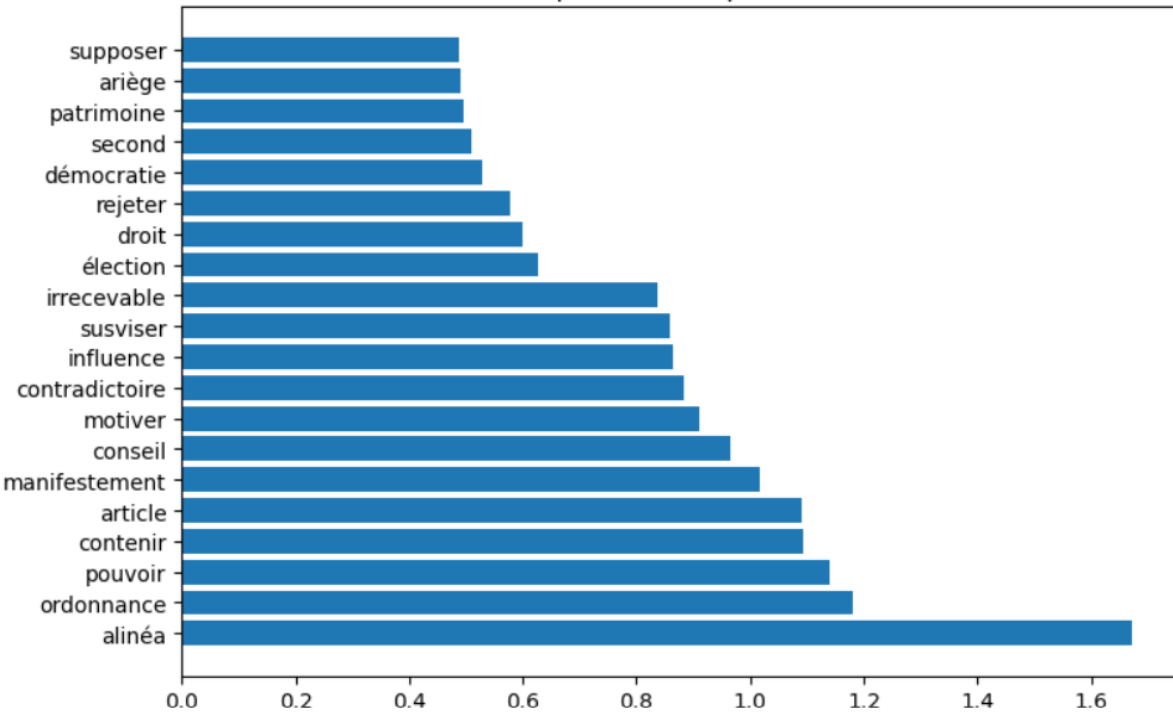
                 self.n_jobs = n_jobs
                 self.remove_saut_ligne = remove_saut_ligne
                 self.remove_decide = remove_decide
                 # self.accented_char = accented_char
                 self.remove_special_char = remove_special_char
                 self.remove_digits = remove_digits
                 self.remove_punctuation = remove_punctuation
                 self.remove_stopword = remove_stopword
                 self.remove_people = remove_people
                 self.remove_place = remove_place
                 self.text_lematization = text_lematization
```

- Amélioration des performances du modèle
- Tester de suppression du texte de loi sur l'article 38 alinéa 2

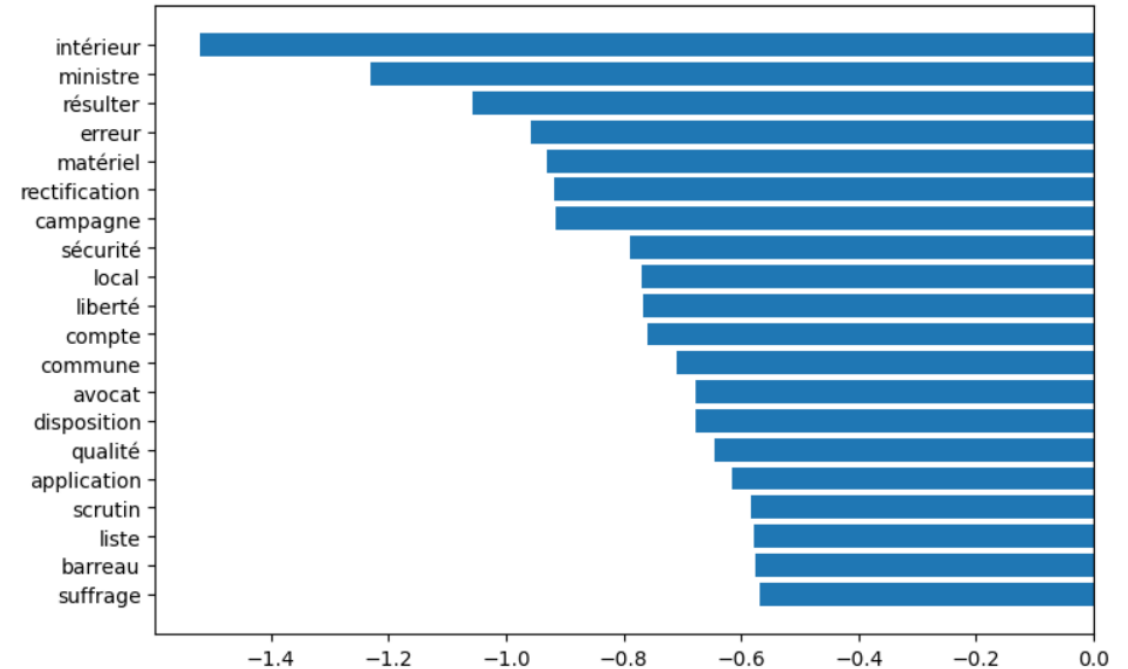
TEST DU MODÈLE AVEC 38 ALINÉA 2 EN APPRENTISSAGE

- Utilisation des Supports Vecteur Machines (SVM) linéaires avec la mention de l'article 38 alinéa 2

Top 20 mots importants



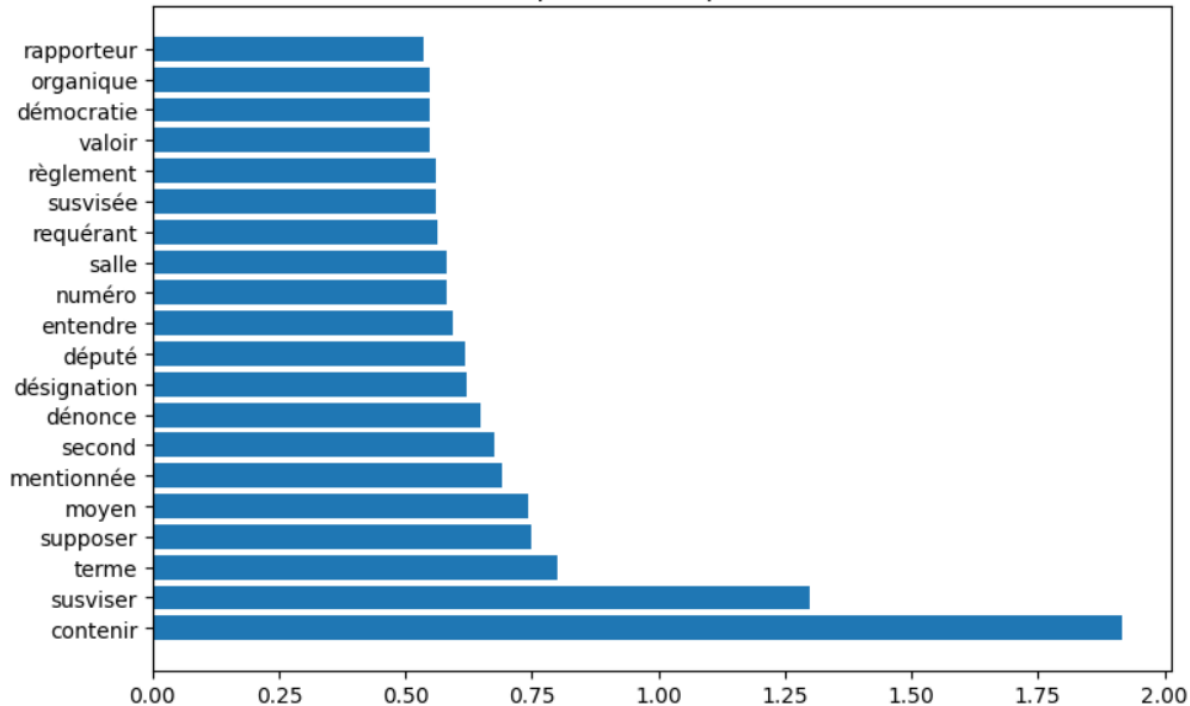
Top 20 mots peu importants



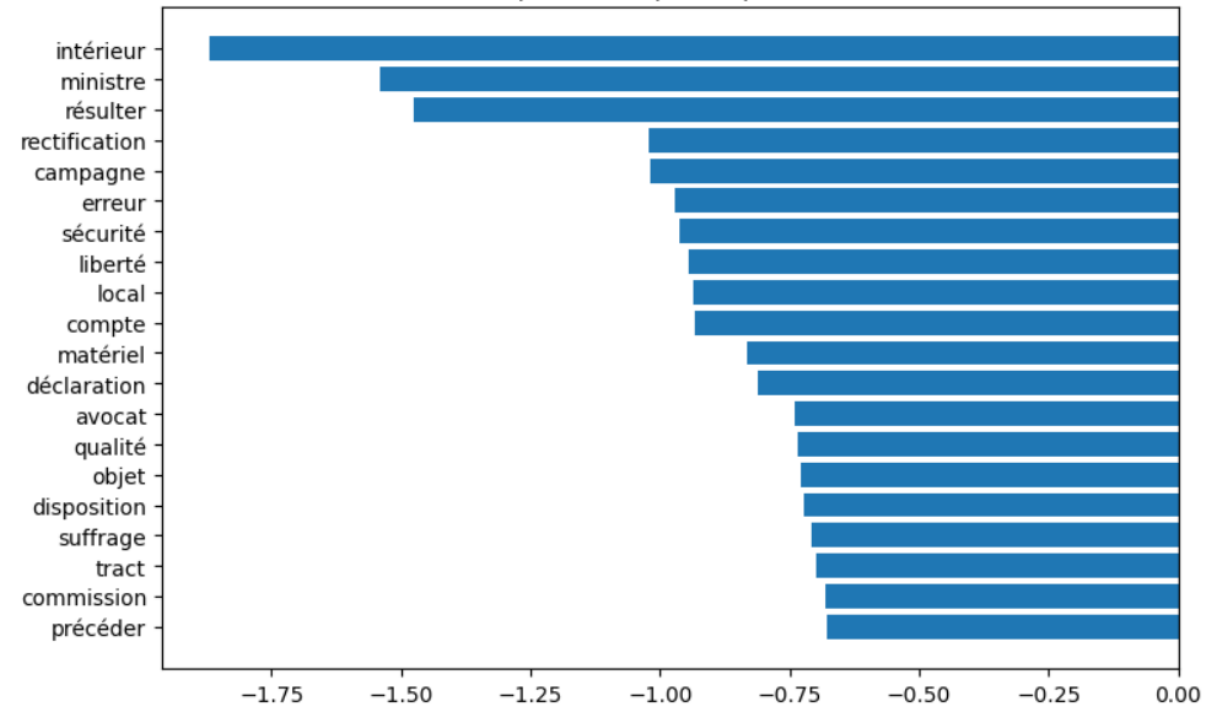
TEST DU MODÈLE SANS 38 ALINÉA 2 EN APPRENTISSAGE

- Utilisation des Supports Vecteur Machines (SVM) linéaires

Top 20 mots importants



Top 20 mots peu importants



Résultats d'algorithmes de TAL (S. Ibghi)

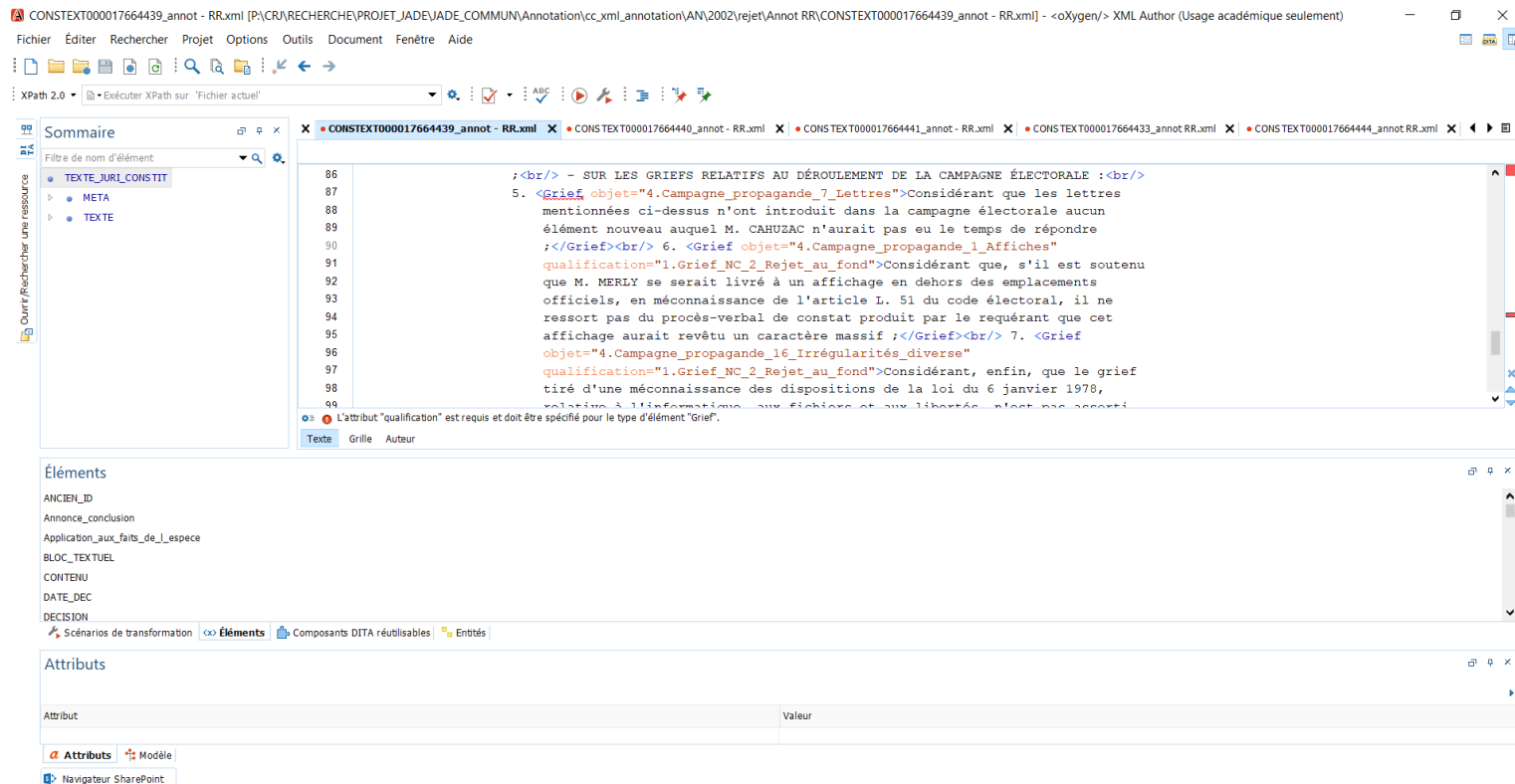
- « *Nous avons essayé une troisième approche, c'est l'extraction de caractéristiques sur la base du modèle entraîné (via ici Camembert) Par la suite nous les utilisons sur des modèles de classification tels que le SVM, la régression logistique, pour les alimenter et faire des prédictions sur les données cibles. Voici un exemple de prédiction à partir de la régression logistique avec une pénalisation de norme L1* »

pred \ inter	Non	Oui	Peut-être	+	++	Partiel. 38 al 2	Peut-être partiel. 38 al 2
Non	719	0	16	3	0	8	43
Oui	31	10	35	104	57	0	0

- **Les cas discordants ont fait l'objet d'un « dialogue » homme / machine**
- **Problème d'explicabilité : rapport aux principes fondamentaux du droit ?**

Etape 2 : Enrichissement des données – l’annotation à des fins de TAL et de statistiques

- 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:

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

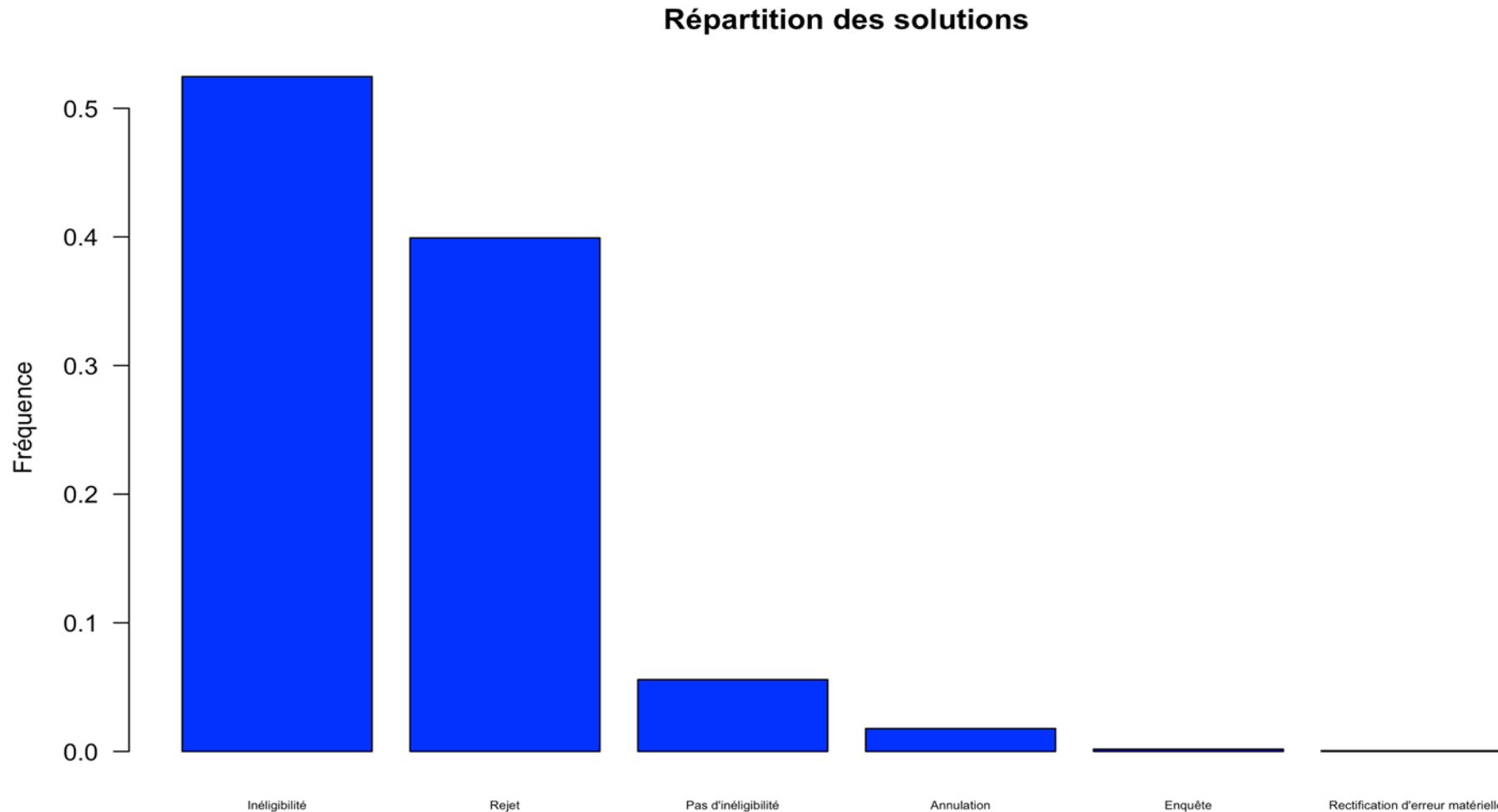
The interface includes a left sidebar with a tree view showing the document structure (e.g., META, TEXTE) and a bottom panel with a table for attributes (Attributs) and a status bar at the bottom.

Etape 3 : Exploitation des données / consultation

- Problèmes des exigences de la recherche en Opendata / Science Ouverte
- Projet de construction d'une interface de consultation de la base des données
- **Problématique CNIL – Données de recherche et données d'intérêt général vs. Données sensibles (pas de réponse de la CNIL depuis septembre)**
 - Point e du § 2 de l'article 6 du RGPD : recherche mission d'intérêt public
 - Articles 9 RGPD et 80 loi 78 : expression universitaire
 - Données politiques = données d'intérêt général ?

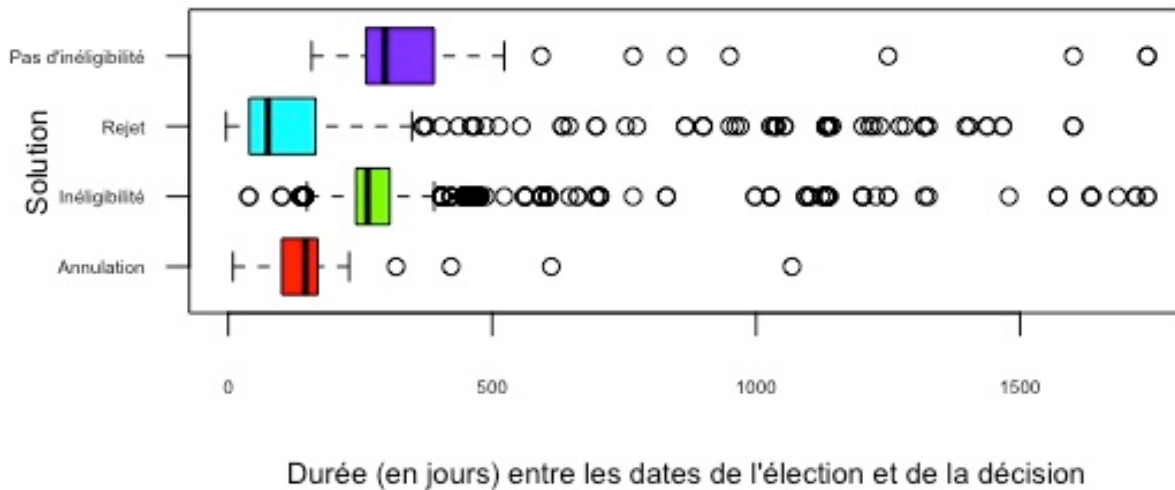
Etape 3 : Exploitation des données / description

- Techniques statistiques descriptives sur le rapport Décisions / Elections

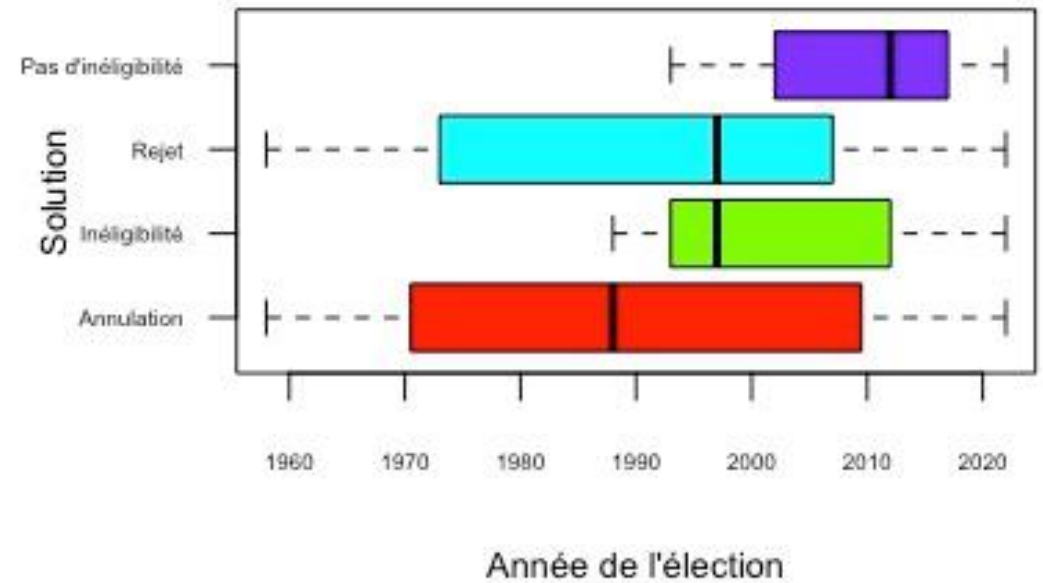


La temporalité du contentieux électoral : : des statistiques confirmatives

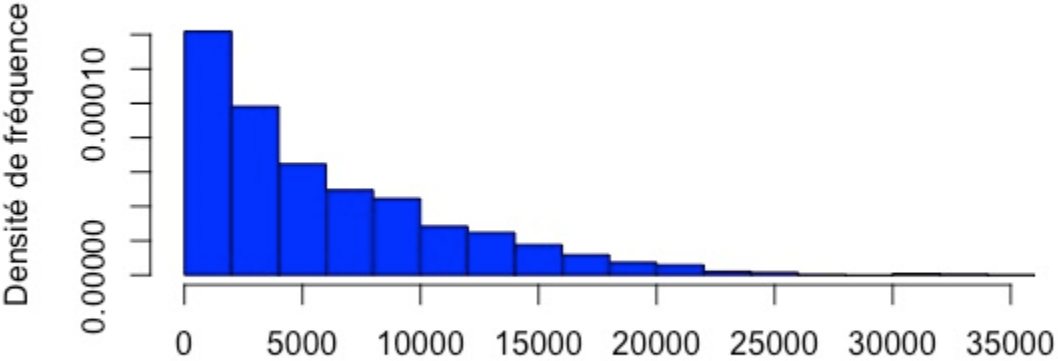
Répartition des durées par solution



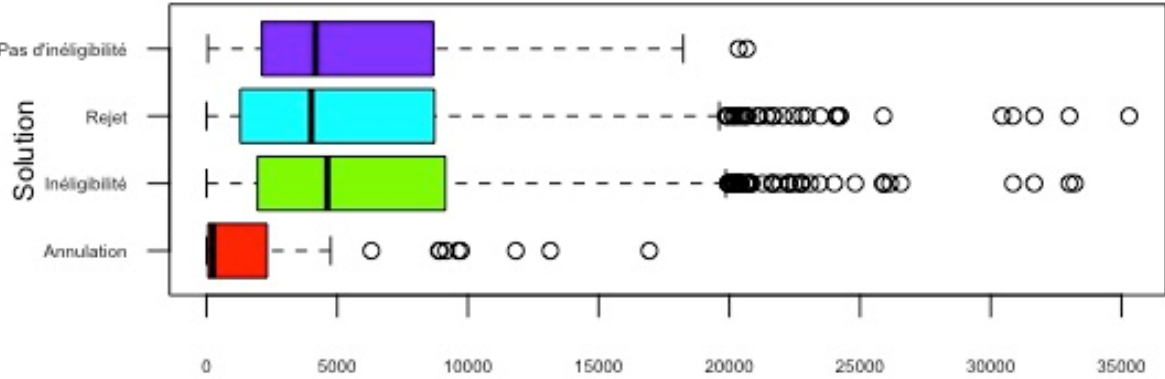
Répartition des années d'élection par solution



Ecart de voix : des statistiques confirmatives

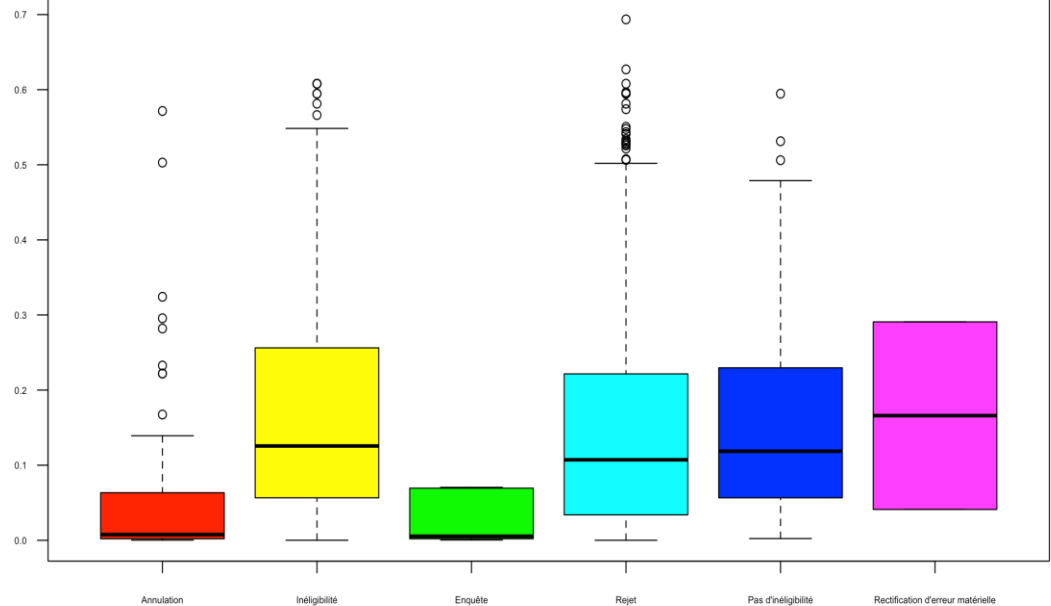


Répartition des écarts de voix entre les candidats 1 et 2 par solution

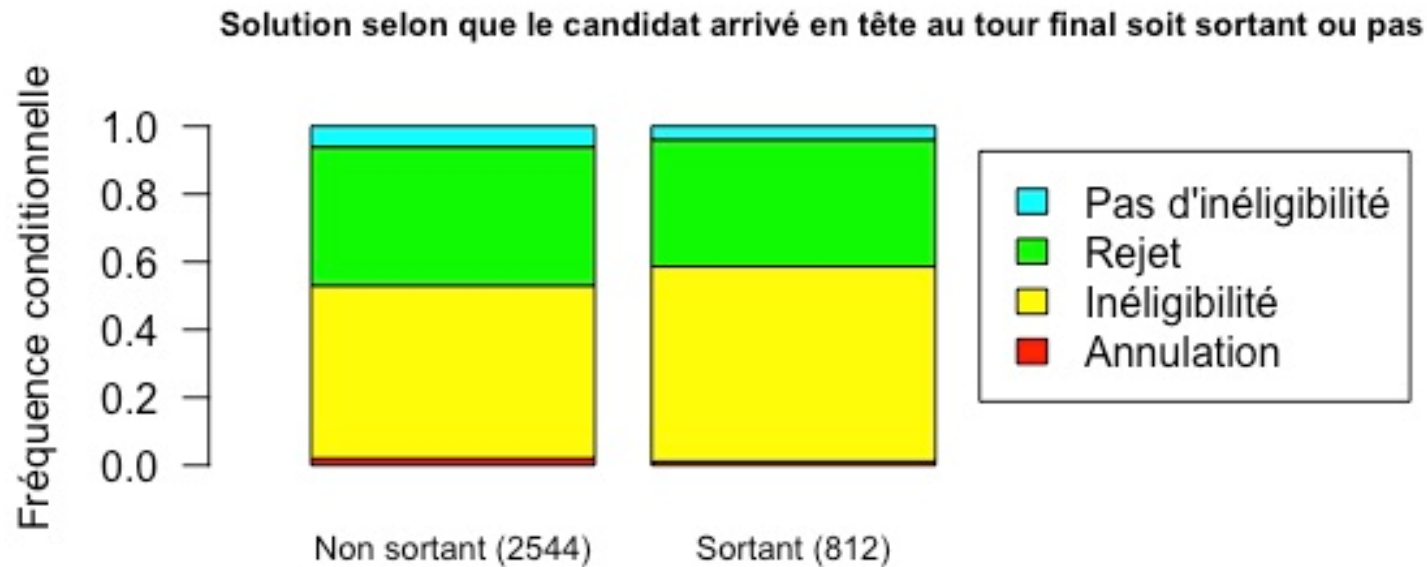


Ecart de voix entre les candidats 1 et 2

Ecart de voix relatifs entre les candidats 1 et 2 selon la solution



Un exemple de résultat nouveau ? La solution contentieuse selon la qualité ou non de sortant...



Etape 3 : Exploitation des données - prévision

- L'hypothèse d'un logiciel d'aide à la consultation / décision pour le citoyen, l'avocat ou le juge en *open data*

Possibilité (Conseil d'Etat) : si ces dispositions art. 120 de la loi de 1978] font obstacle à ce qu'une juridiction fonde sa décision sur les seuls résultats d'un traitement automatisé d'informations, elles n'ont en revanche ni pour objet ni pour effet de lui interdire de prendre en compte, parmi d'autres éléments d'appréciation, les résultats d'un tel traitement ; qu'ainsi, en jugeant, pour accueillir l'appel de M. X, que les dispositions précitées s'opposaient à ce que les calculs statistiques automatisés invoqués par la caisse soient pris en compte pour apprécier le comportement professionnel de l'intéressé, la section des assurances sociales du conseil national de l'ordre des médecins a commis une erreur de droit » (CE, 4 fev. 2004, n°240023, *CPAM de la Gironde*).

Limite : L. 10 CJA – interdiction d'utiliser ces système sur les noms des juges

- Techniques statistiques « prédictives » : système de variables déterminantes (griefs, écarts de voix, partis politiques, etc.) ?
- Systèmes de clustering par « machine learning » / arbres de décision ?
- IA génératives ?

Conclusion - Difficultés et Perspectives : le projet JADE et la Justice algorithmique

- Réussite à mener à bien le projet concernant le Conseil constitutionnel ?
Très forts besoins humains, difficultés de l'interdisciplinarité
- Quels résultats pour le Conseil constitutionnel ?
- Quid de la justice administrative, Conseil d'Etat et TA ?
- Quel impact sur la jurisprudence ?
- Quel impact sur la théorie du droit ?
- Quel succès du projet JADE UGA/UCA dans le cadre de l'ANR ?
- Le projet JADE, un projet pilote ? Quels autres contentieux ? Quelles problématiques ?