

Présentation des ressources

Jeux de données, API et outils IA utiles pour les défis

Emmanuel Raviart & Pierre Drege
3 et 4 juillet 2026

Données brutes

Open data de l'Assemblée nationale

- <https://data.assemblee-nationale.fr/>
- Députés, organes, amendements, votes, réunions, comptes rendus, dossiers législatifs, questions
- Formats XML, JSON (et CSV)
- Manque les textes (projets et propositions de lois, rapports, etc)

Open data du Sénat

- <https://data.senat.fr/>
- Sénateurs, organes, textes, amendements, comptes rendus, dossiers législatifs, questions
- Format : dump PostgreSQL ancien

Open data de Légifrance

Direction de l'information légale et administrative (Dila)

- <https://echanges.dila.gouv.fr/OPENDATA/>
- Journaux officiels, textes, sections et articles
- Textes consolidés (codes, etc)
- Format : XML

Carte des données utiles

Quelques familles disponibles

- Vie parlementaire
 - Députés, sénateurs, mandats, groupes, organes
 - Réunions, commissions, séances publiques
 - Votes, scrutins, questions
- Production normative
 - Projets et propositions de loi
 - Amendements, rapports, dossiers législatifs
 - Lois, codes, décrets, arrêtés, Journal officiel

Open data de Service-public.fr

Direction de l'information légale et administrative (Dila)

- <https://echanges.dila.gouv.fr/OPENDATA/>
- Guide des droits et démarches (particuliers, associations et entreprises)
 - Fiches pratiques de Service-public.fr
 - Parcours administratifs, publics concernés, conditions
 - Renvois vers textes, formulaires et démarches en ligne
- Annuaire de l'administration
- Format : XML

Nettoyage et archivage

Données nettoyées de l'Assemblée nationale

- <https://git.en-root.org/tricoteuses/data/assemblee-nettoye>
- Données nettoyées
- Converties en JSON
- Mises sous Git

Données nettoyées du Sénat

- <https://git.tricoteuses.fr/senat>
- Conversion de PostgreSQL en JSON
- Correction de l'encodage des caractères
- Mises sous Git

Données de Légifrance et Service-public

- <https://git.tricoteuses.fr/dila>
- Mises sous Git
- XML
- Et Markdown (HTML)
- Et textes consolidés :
 - <https://git.tricoteuses.fr/codes>
 - https://git.tricoteuses.fr/constitution/constitution_du_4_octobre_1958

Documentation

Documentation

- Documentation *JSON Schema* :
 - <https://www.tricoteuses.fr/assemblee/schemas>
 - <https://www.tricoteuses.fr/senat/schemas>
 - <https://www.tricoteuses.fr/legifrance/schemas>
- Documentation insérée dans les données

Base de données

Base de données PostgreSQL

- Base de données unifiée *Canutes*
- Dumps hebdomadaire : <https://dump.tricoteuses.fr/>
- Accès direct :
 - `hackathon2026.leximpact.dev`
 - Port : 5432
 - Utilisateur : `hackathon2026`
 - Mot de passe : `hotel-de-lassay`

Données moissonnées

Données ajoutées à l'open data

- Textes de l'Assemblée
- Données temps réel
 - Informations extraites des vidéos
 - Dérouleur de la séance publique de l'Assemblée (*jaune*)
 - Éliasse (suivi en temps réel des amendements)
 - Scrutins
- Ajoutées à la base de données canutes

API

Deux API

- API Parlement
 - Données à plat dans des tables
 - Données de l'Assemblée et du Sénat uniquement
 - <https://parlement.tricoteuses.fr/docs>
- API PostgREST :
 - <https://db.code4code.eu/canutes/>
 - Données structurées
 - Tout le contenu de la base *Canutes*

Model Context Protocol (MCP)

Premier MCP

- MCP Parlement
 - <https://www.tricoteuses.fr/services/mcp-parlement>
 - Données de l'Assemblée et du Sénat uniquement
 - MCP classique : outils correspondent peu ou prou à Tricoteuses API Parlement

Deuxième MCP

- MCP Moulineuse
 - <https://www.tricoteuses.fr/services/mcp-moulineuse>
ou <https://mcp.hackathon2026.leximpact.dev/mcp>
 - Tout le contenu de la base *Canutes*
 - Outils bas niveaux : SQL, scripts JavaScript, etc
 - Recettes métiers

MCP comme API

- MCP Moulineuse = API survitaminée
- API permettant d'exécuter du SQL et du JavaScript

En travaux

Contribuer
<https://www.tricoteuses.fr/>

