

# Pour des IA Francophones



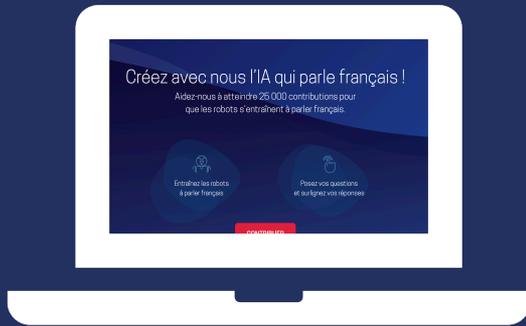
# Programme

Rétrospective (3min)

Travaux en cours (10min)

Nos problèmes, vos idées  
#discussion (10min)

# Rétrospective :



*Open-source*



# Rétrospective :

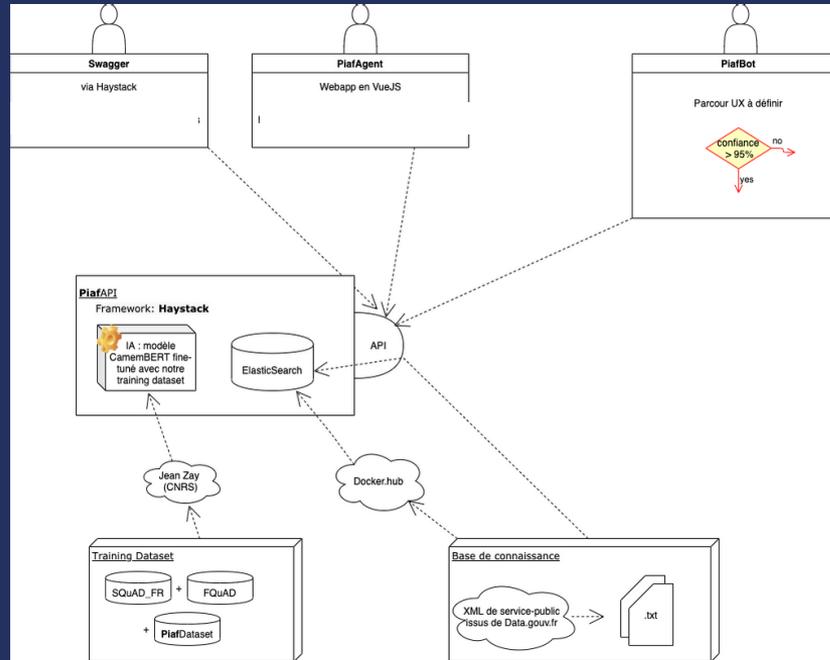
- Micro-travail VS contribution engagée => 🥵
- Ressource native VS multilingue =>

# Travaux en cours :

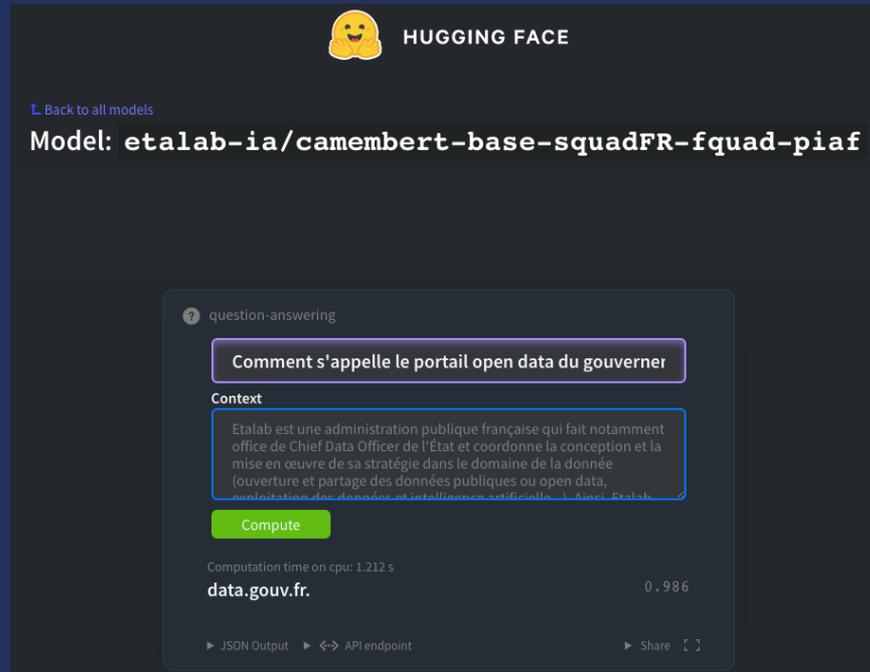
Passer du dataset, à une application clé en main



# Travaux en cours :



# Travaux en cours : Modèle entraîné



The screenshot shows the Hugging Face interface for a model. At the top, there is the Hugging Face logo and the text 'HUGGING FACE'. Below that, there is a link 'Back to all models' and the model name 'Model: etalab-ia/camembert-base-squadFR-fquad-piaf'. The main content area shows a question-answering task. The question is 'Comment s'appelle le portail open data du gouverner'. The context provided is 'Etalab est une administration publique française qui fait notamment office de Chief Data Officer de l'État et coordonne la conception et la mise en œuvre de sa stratégie dans le domaine de la donnée (ouverture et partage des données publiques ou open data, exploitation des données et intelligence artificielle...)'. The model's answer is 'data.gouv.fr.'. Below the answer, there is a 'Compute' button, the computation time '1.212 s', and the score '0.986'. At the bottom, there are links for 'JSON Output', 'API endpoint', and 'Share'.

HUGGING FACE

[Back to all models](#)

Model: **etalab-ia/camembert-base-squadFR-fquad-piaf**

question-answering

Comment s'appelle le portail open data du gouverner

Context

Etalab est une administration publique française qui fait notamment office de Chief Data Officer de l'État et coordonne la conception et la mise en œuvre de sa stratégie dans le domaine de la donnée (ouverture et partage des données publiques ou open data, exploitation des données et intelligence artificielle...). Ainsi, Etalab

Compute

Computation time on cpu: 1.212 s

**data.gouv.fr.** 0.986

[JSON Output](#) [API endpoint](#) [Share](#)

<https://huggingface.co/etalab-ia/camembert-base-squadFR-fquad-piaf>

# Travaux en cours : PiafAgent



Piaf & Service-public.fr

Une autre façon de trouver une information

[Chercher](#)



PIAFAgent & service-public.fr

**Indice de confiance : 84 %**

Congé maternité **d'une salariée du secteur privé** : Vous bénéficiez d'un congé de maternité durant la période qui se situe autour de la date présumée de votre accouchement. Sa durée est variable, en fonction du nombre d'enfants à naître ou déjà à charge. Il comporte une période de congé prénatal et un congé postnatal. Vous bénéficiez d'une indemnisation versée par la Sécurité sociale. Qui est concerné ? : Vous bénéficiez automatiquement d'un congé de maternité, en partie avant votre accouchement et en partie après. Le congé de maternité ...

[Lien vers la fiche](#)

**Indice de confiance : 64 %**

Congé maternité **d'une salariée du secteur privé** : Vous bénéficiez d'un congé de maternité durant la période qui se situe autour de la date présumée de votre accouchement. Sa durée est variable, en fonction du nombre d'enfants à naître ou déjà à charge. Il comporte une période de congé prénatal et un congé postnatal. Vous bénéficiez d'une indemnisation versée par la Sécurité sociale. Qui est concerné ? : Vous bénéficiez automatiquement d'un congé de maternité, en partie avant votre accouchement et en partie après. Le congé de maternité ...

[Lien vers la fiche](#)

[Poser une nouvelle question](#)

# Travaux en cours : PiafBot

The image shows a screenshot of the PiafBot chatbot interface on the service-public.fr website. The interface is split into two main sections: a header/sidebar on the left and a chat window on the right.

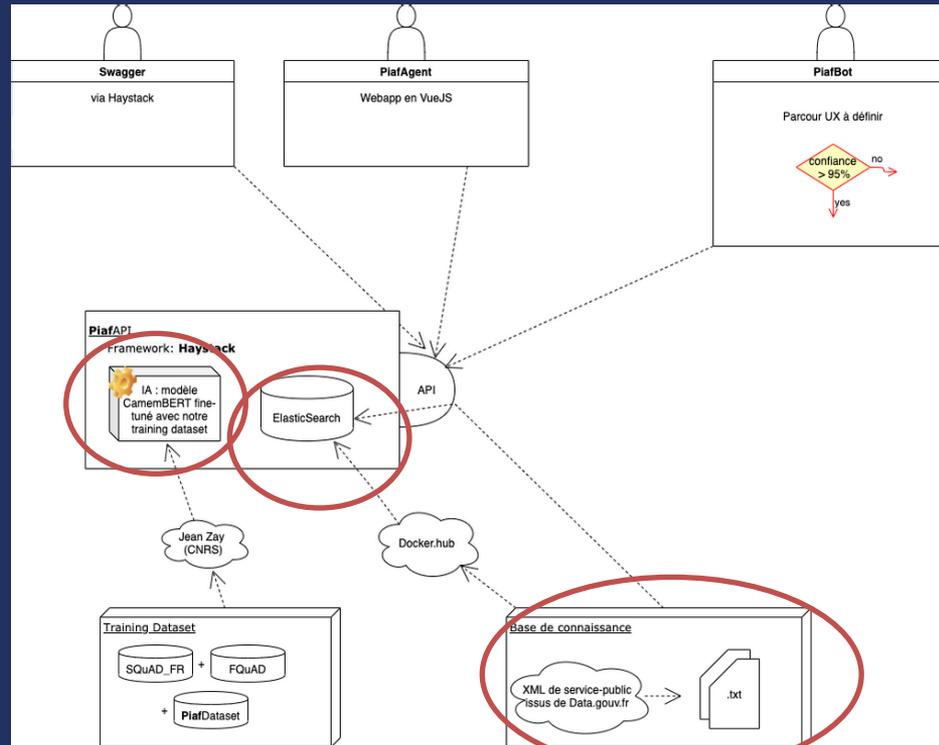
**Header/Sidebar (Left):**

- Logo of the French Government: **GOVERNEMENT** with the motto "États Généraux Français".
- Service Public logo: A stylized head profile with a red and blue color scheme.
- Text: **Service Public**
- Accessibility toggle: A switch labeled "Accessibilité".
- Footer: "Conçu et développé avec ❤️ par la **Fabrique numérique**".

**Chat Window (Right):**

- Bot Message 1:** "Bonjour, je suis le prototype de chatbot de service-public.fr 🙋"
- Bot Message 2:** "Toutes mes réponses sont automatisées. Posez-moi directement une question ou sélectionnez un des thèmes ci-dessous :"
- Thème Selections:** A grid of buttons with icons for: "Papiers - Citoyenneté", "Famille", "Social - Santé", "Travail", "Logement", "Transport", "Argent", "Justice", "Etranger", and "Loisirs".
- User Message:** "Quel est le coût d'un passeport ?"
- Bot Message 3:** "J'ai trouvé la réponse suivante: 86 €"
- Bot Message 4:** "Dans le corps de texte suivant: Passeport d'un majeur : première demande : Accessible avec un compte ANTS ou via France Connect"
- Bot Message 5:** "En France:"
- Bot Message 6:** "Coût: as général Le passeport coûte 86 €"
- Bot Message 7:** "Se munir d'une carte bancaire"
- Input Field:** "Posez votre question ..."

# Nos problèmes :



# Nos pistes actuelles :

1. Mesure quanti : **BM25 VS Embeddings**
2. Voir si **d'autres modèles de SBERT** sont plus performants
3. Base de connaissance => que faire pour aider le Reader. Voir si Haystack permet de demander au retriever de chercher dans un texte qui contient les titres et sous-titre, tandis que le Reader ne cherche que dans le texte pur (sans le bruit apporté par les meta infos)
4. Mise en place des **filtres**, comme des # par thématique
5. Essayer le **DPR**
  - Est qu'on fait une expérience en anglais d'abord ?
  - On demande a Timo de Haystack ? Djamé.S ? J.Chaumond ? code du travail numérique ?
  - Regarder ce que ça coûte d'entraîner un DPR en français
6. Reader : Utiliser les 25k / fiches de SPF pour **fine-tuner CamemBERT** ?
7. **Mixer plusieurs retrievers**

**Questions & Idées ?**

**Des nouvelles applications ?**