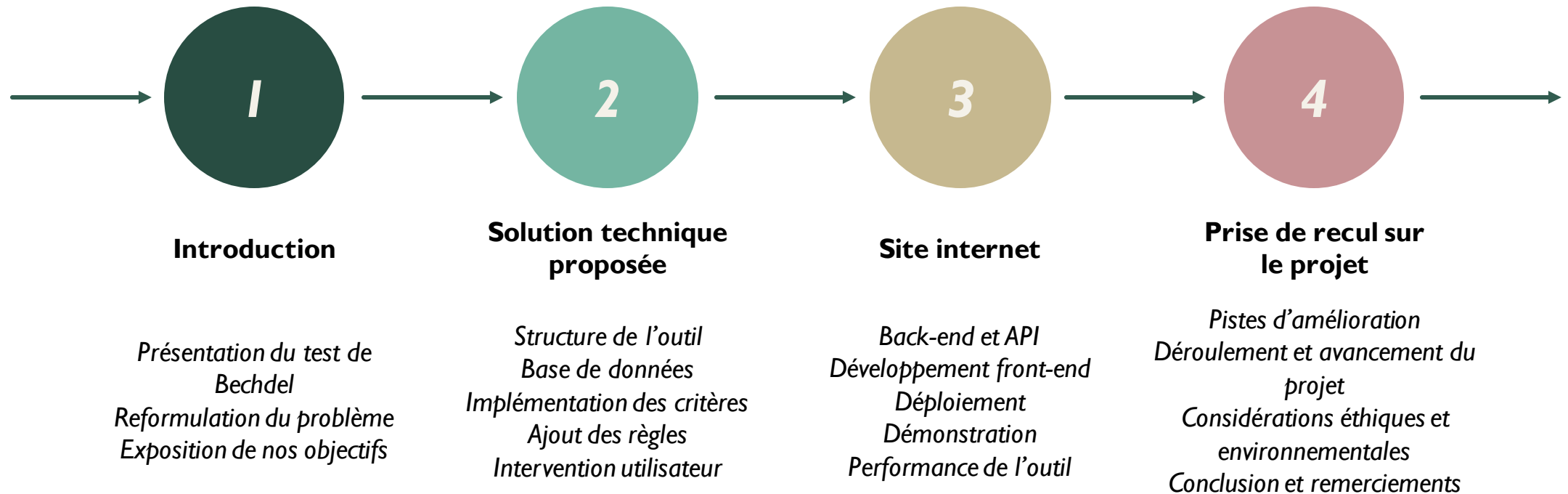




AUTOMATISATION DU TEST DE BECHDEL

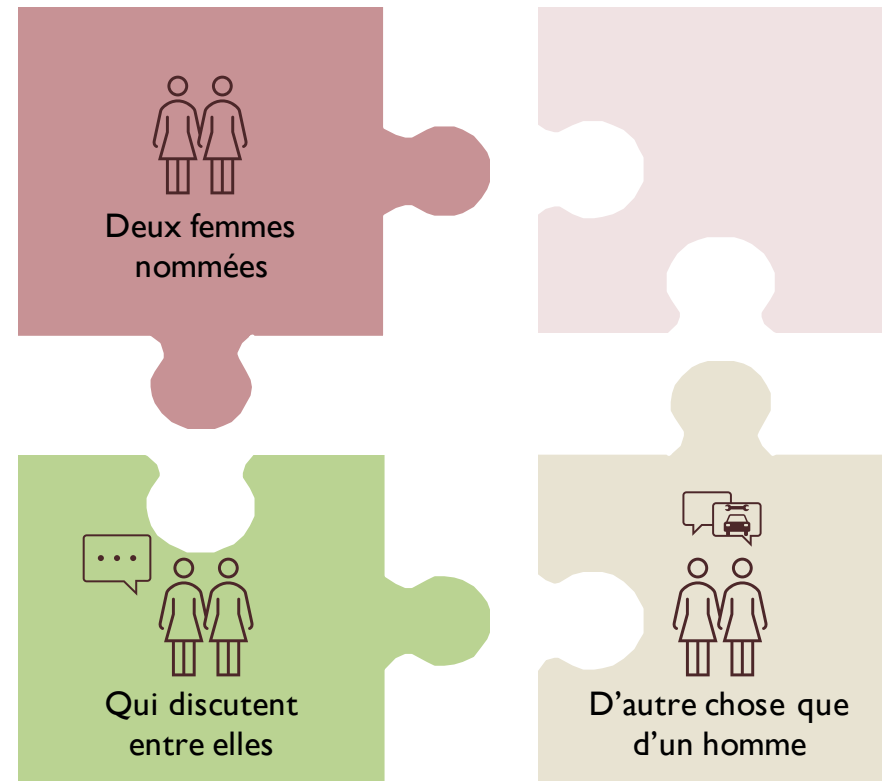
SOUTENANCE FINALE – PROJET INFONUM

PLAN



PRÉSENTATION DU TEST DE BECHDEL

On dit qu'un film valide le test de Bechdel s'il contient :



REFORMULATION DU PROBLÈME, NOS OBJECTIFS

OBJECTIF

*Créer un outil performant, rapide et accessible
pour étudier si un film valide le test de Bechdel*

Structure proposée

Création d'un site en ligne

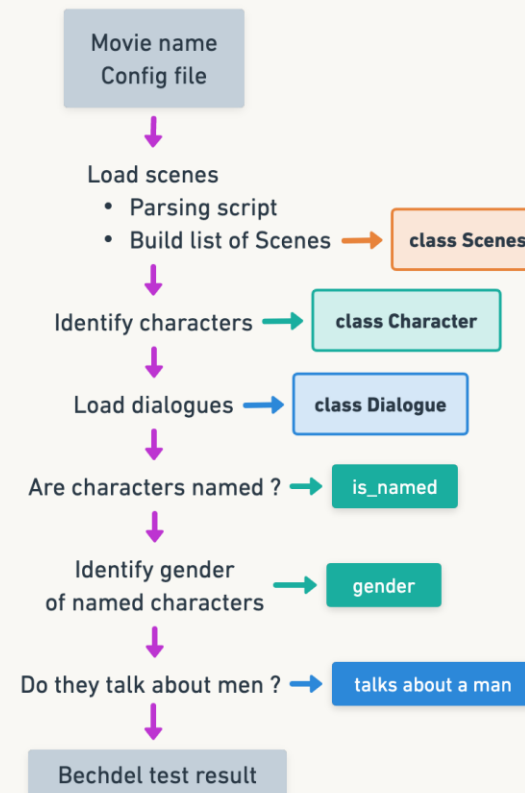
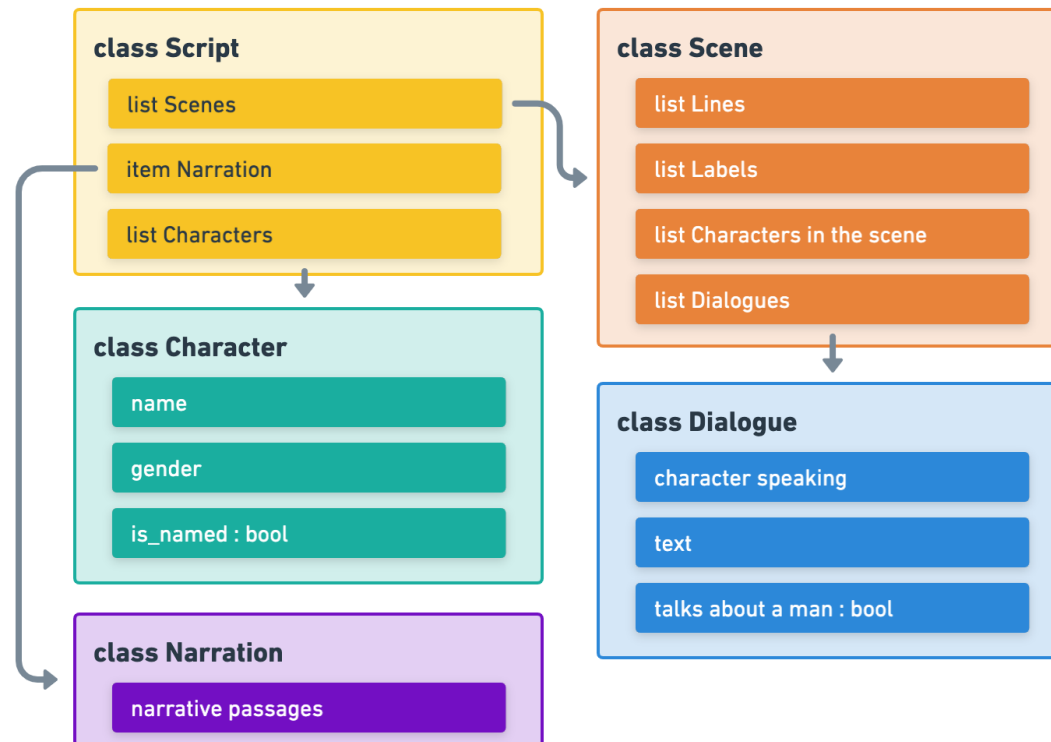
&

Contribution Open Source

Qui prend en entrée un Script
de film



STRUCTURE DE L'OUTIL

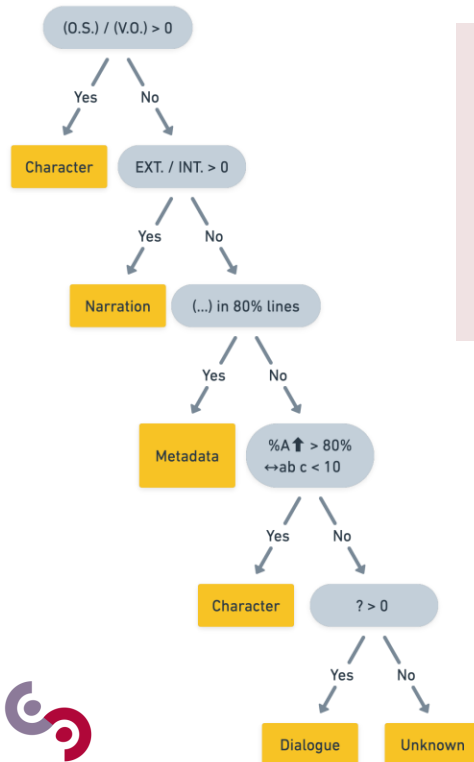


BASE DE DONNÉES D'ENTRAÎNEMENT



PARSING DES SCRIPTS

Méthode naïve

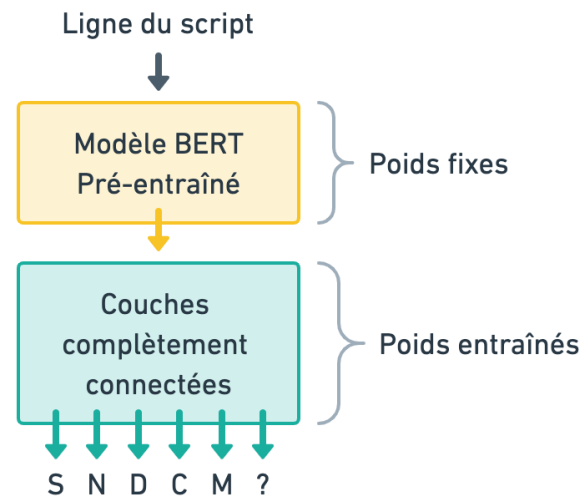


Arbre de décision
construit manuellement

Parse correctement les
2/3 du dataset



Méthode Deep Learning



Modèle entraîné sur les
labels inférés par l'arbre

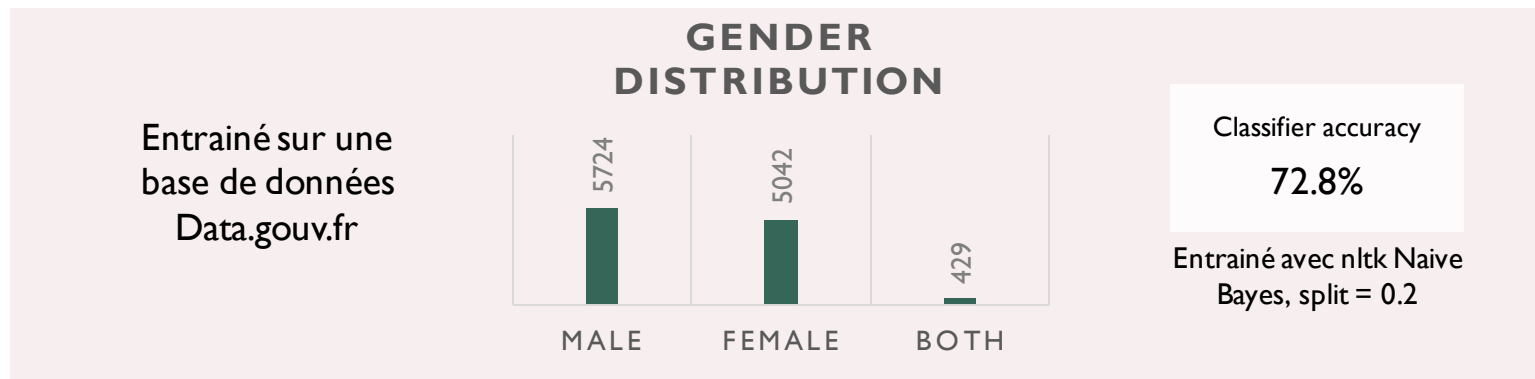
Parse correctement le
reste du dataset

Accuracy de 85%



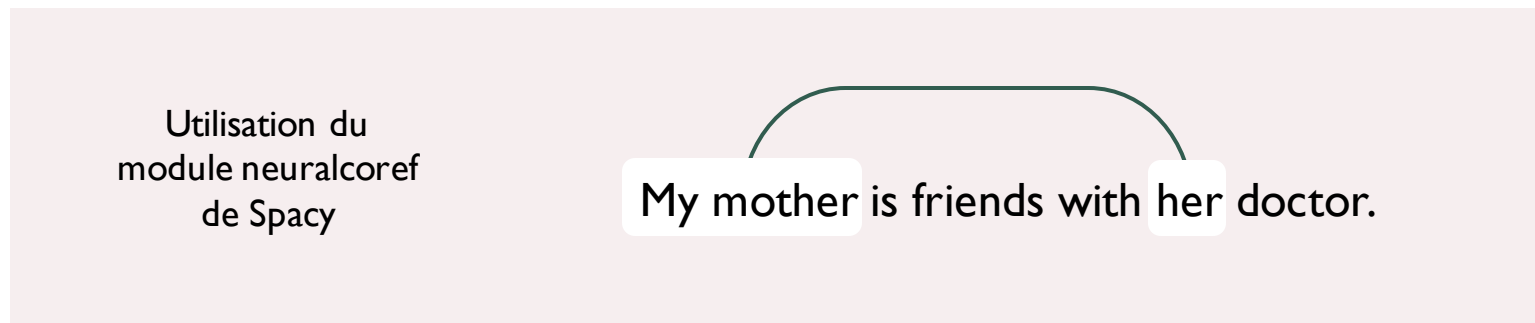
PRÉDICTION DU GENRE DES PERSONNAGES

Méthode 1 : classification sur les prénoms



Peu performant dans les faits, sur des prénoms fictifs

Méthode 2 : Coréférence neuronale sur les passages narratifs



TOPIC MODELING SUR LE CONTENU DES CONVERSATIONS

Une méthode naïve

JUDY

Uh, **Bogo** Otterton? Been coming to your yoga class for like 6 years? The little **boy**, you know ?

Score 2/3 ✓

NANGI

I have no memory of this beaver.

Détection de mots clés dans les répliques.

JUDY

He's an otter actually.

En pratique, pas de contre exemples trouvés dans nos scripts

Des éventuelles améliorations

NANGI

Do you know Jack? He's the baker accross the street.

Score 3/3 ✗

JUDY

Yeah, a very good baker! Oh look, that's **Barack Obama** leaving the bakery!

Ajout des prénoms communs aux mots clés ?

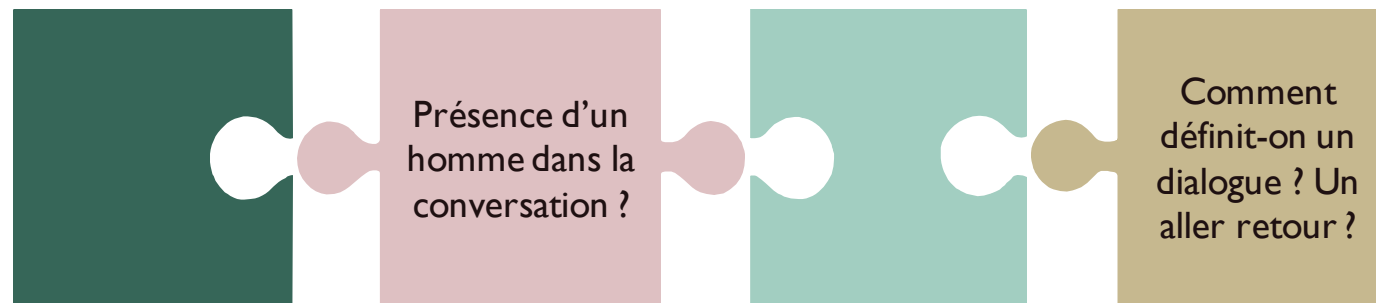
NANGI

They must have talked a little.

"Baker, they" : Ajout de la coréférence dans les répliques



RÉFLEXION SUR LA DURETÉ DES RÈGLES



Ajout des paramètres



Uniquement des femmes dans la scène



Intégralité du dialogue pas à propos d'un homme

Que 3 configurations possibles :
(V,V), (V,F), (F,F)



AJOUT INTERVENTION UTILISATEUR

LIST OF CHARACTERS

NAME	Gender prediction ▼	NAME	Gender prediction ▼
	Male Female Non-binary		
<p>TEST AGAIN !</p>			



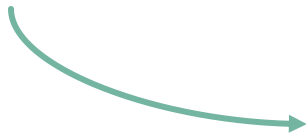
Capacité donnée à
l'utilisateur de corriger
la prédiction de genre



BACK-END ET API

Back-end

 FastAPI



POST

Upload d'un fichier texte,
conversion en objet Script,
calcul du score



Re-calcul du score après une
correction de genre par
l'utilisateur-ice



GET

Récupération du score étant
donné un identifiant de script



Récupération du contenu des
scènes clé étant donné un
identifiant de script



Base de données
d'objets scripts

DÉVELOPPEMENT FRONT-END

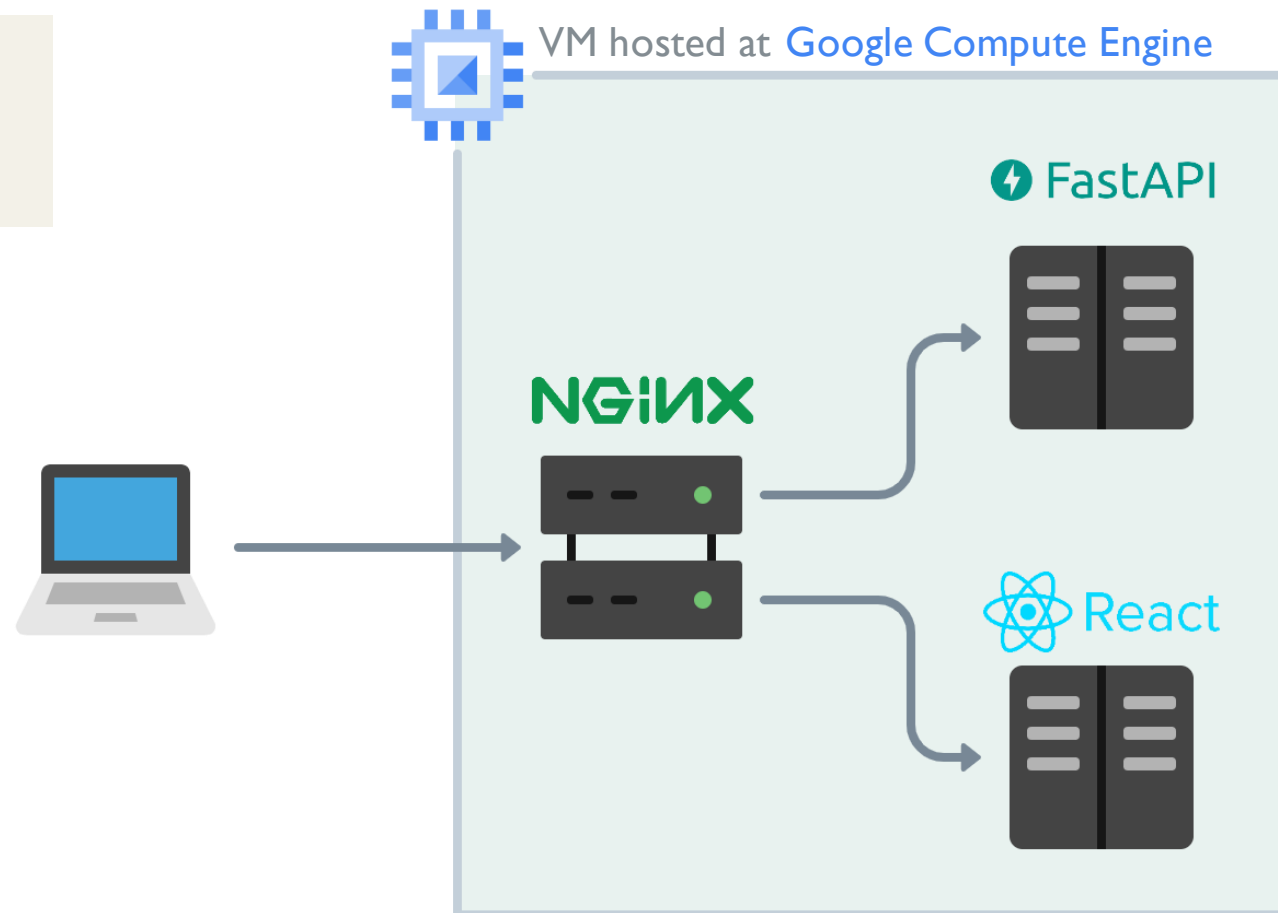
The screenshot shows a web application interface for a Tetris game. The interface is divided into several sections:

- Welcome !**: A section with a placeholder for a welcome message.
- Upload a script**: A section with a placeholder for a script and an **Upload** button.
- Parameters**: A section with two toggle switches and a placeholder for a parameter.
- Results**: A section displaying the **Score : X/3** and a dropdown menu labeled **Browse the scenes that validate the score :**. Below the dropdown is a Tetris board with a yellow Tetris piece in the center.
- Characters**: A section with two dropdown menus and a **Test again !** button.



DÉPLOIEMENT

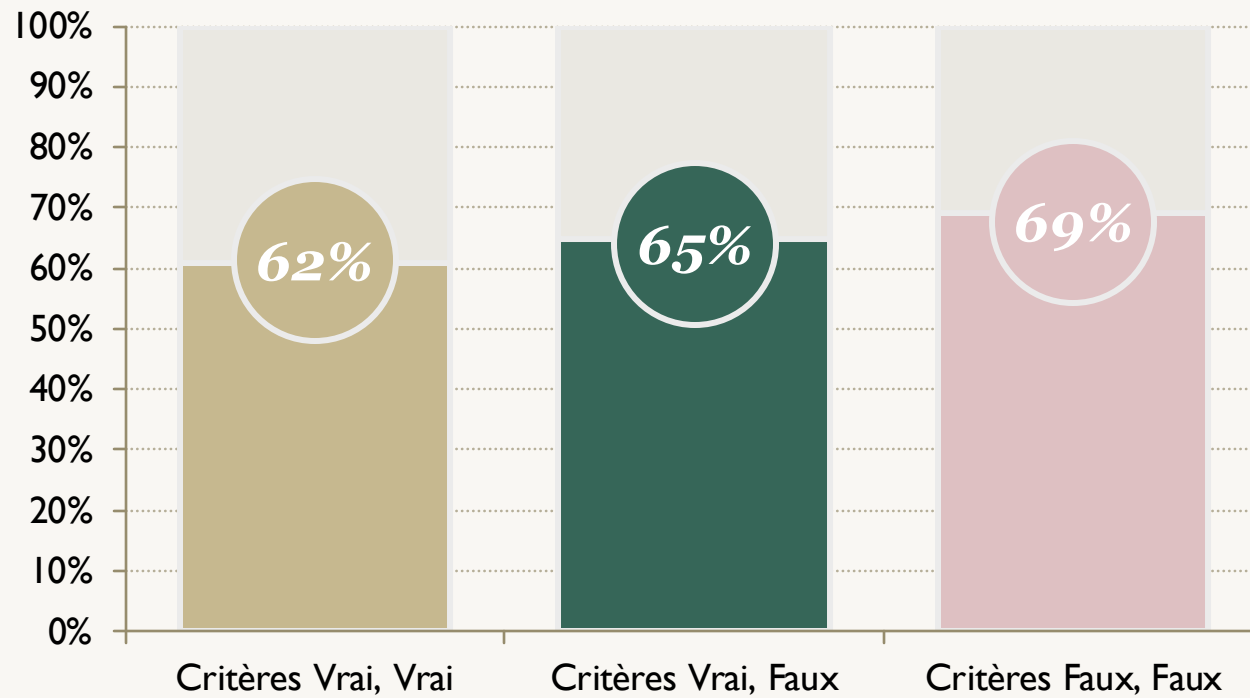
Utilisation de Nginx comme
reverse proxy



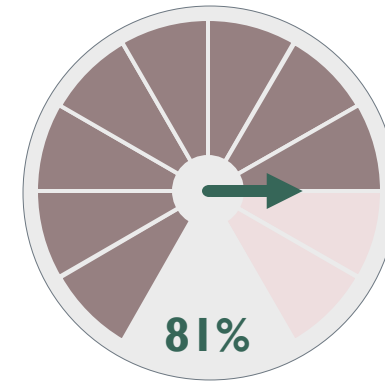
DÉMO DE L'OUTIL

[Lien vers le site internet](#)

PERFORMANCE DE L'OUTIL

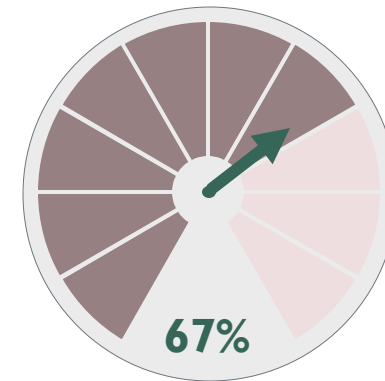


Accuracies



DE PRÉCISION

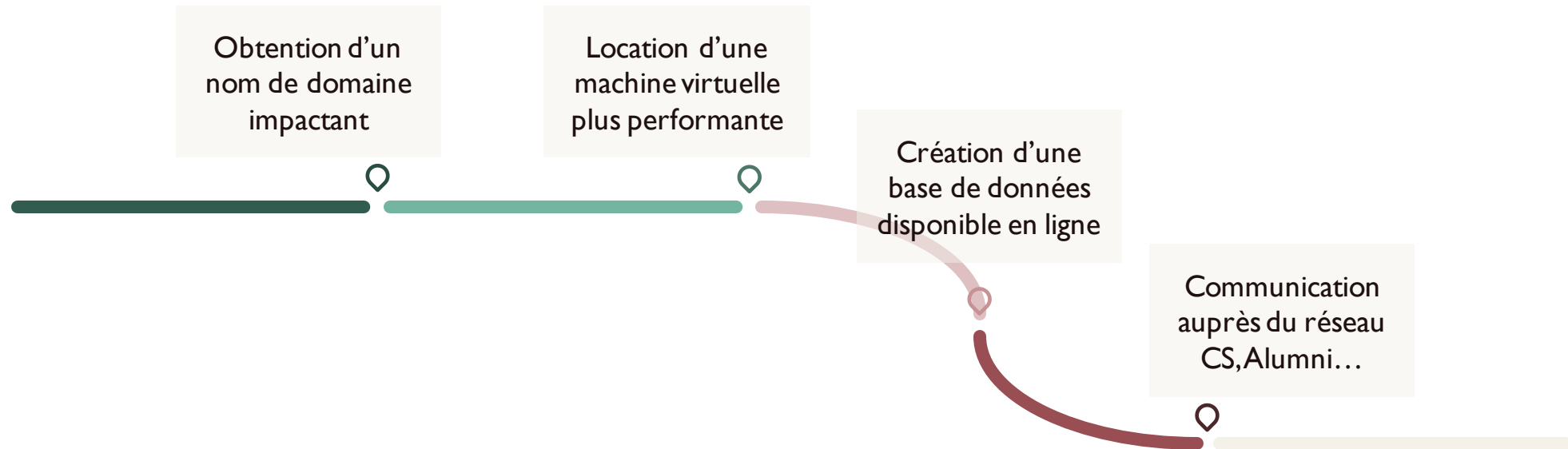
Sur la base de 700 films,
En configuration T, F



DE RAPPEL

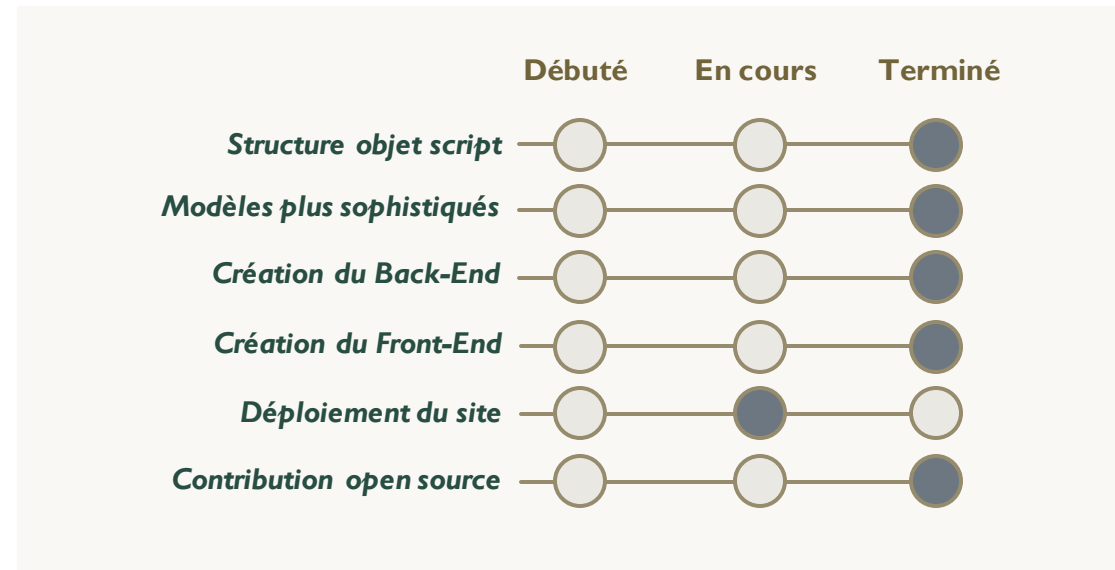
Sur la base de 700 films,
En configuration F, F

PISTES D'AMÉLIORATION



DÉROULEMENT ET AVANCEMENT DU PROJET

Avancement des différentes étapes



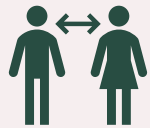
Organisation en groupe



CONSIDÉRATIONS ÉTHIQUES ET ENVIRONNEMENTALES

Enjeux éthiques

Écueil dans la prédiction de genre
d'un personnage



Inclusion des individus non-
binaires dans l'outil (et donc dans
le test)

Impact environnemental

5.9 gCO₂e
Pour entraîner le
modèle de parsing

0.9 gCO₂e
Pour calculer le
score d'un film

CONCLUSION ET REMERCIEMENTS



Performant



Rapide



Accessible

Remerciements

Nous tenons à remercier les encadrants côté CentraleSupélec pour leur accompagnement pendant ce projet, et nous tenons à témoigner toute notre reconnaissance à Théo Rubenach, notre encadrant côté Illuin, pour son temps et son aide précieuse à la réalisation du projet.

MERCI POUR VOTRE
ATTENTION

TEMPS D'ÉCHANGE

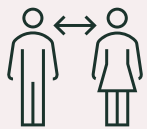


ANNEXES

CONSIDÉRATIONS ÉTHIQUES ET ENVIRONNEMENTALES

Enjeux éthiques

Écueil dans la prédiction de genre
d'un personnage



Inclusion des individus non-
binaires dans l'outil (et donc dans
le test)

Impact environnemental

	Entraînement du modèle de parsing	Inférence de l'outil
Mix électrique intensité carbone	55 gCO ₂ e/kWh	128,94 gCO ₂ e/MJ
Energie consommée (kWh)	0,107	0,0019
Emissions totales (gCO ₂ e)	5,9	0,9