



CARTABLE
FANTASTIQUE



PARIS
DIGITAL **LAB**

Soutenance de fin de projet

22/10/2021

Équipe

- Durant 7 semaines, une équipe de 3 étudiants de CentraleSupélec conçoit et développe un **Minimum Viable Product** dans le cadre du Paris Digital Lab.
- Les étudiants sont encadré technique et méthodologiquement par **Théo Rubenach**, Data Scientist de la société ILLUIN Technology, partenaire IA et Digital du PDL
- Un référent de Centralesupélec coach l'équipe sur les usages : **Renaud Monnet**



Ali CHARARA
ali.charara@student-cs.fr



Benjamin POUX
b.pouxbouret@gmail.com



Charlotte SASSON
charlotte.sasson@student-cs.fr



Renaud MONNET
Co-Directeur de la Chaire
d'innovation Digitale de
CentraleSupélec
CIO CentraleSupélec



Théo RUBENACH
Coach Technique
du Paris Digital Lab
Data Scientist
at ILLUIN Technology



— TABLE DES MATIÈRES

01	Mise en Contexte	05	Interface de correction
02	Traitement des données	06	Taggage automatique
03	Adaptation des exercices	07	Organisation du rendu
04	Rendu Graphique	08	Conclusion

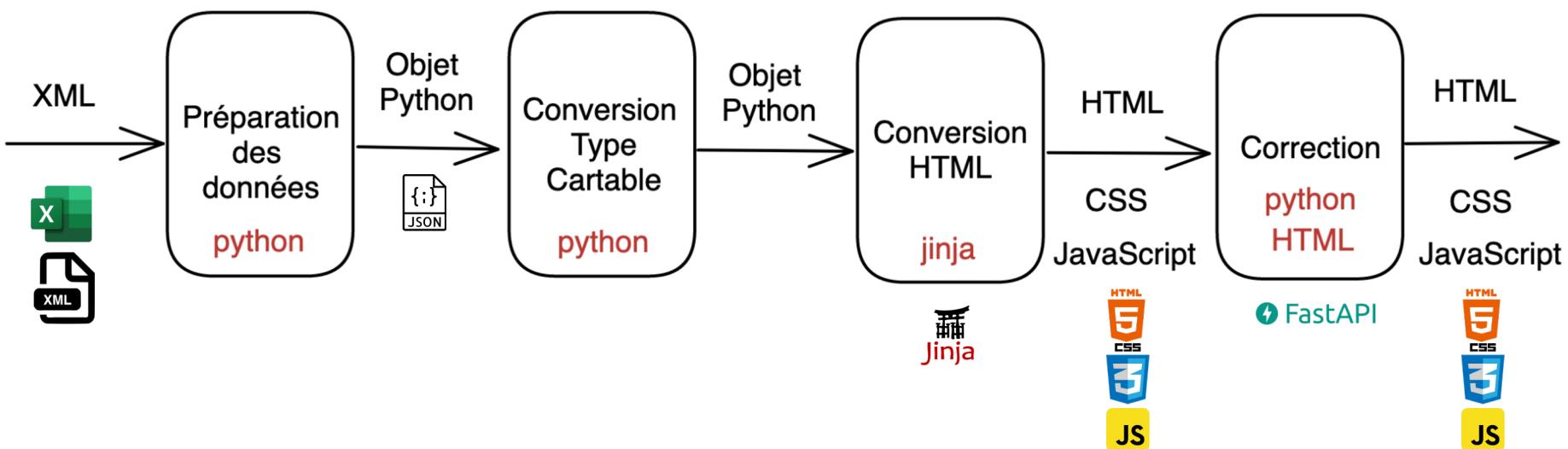
Mise en contexte

Le cartable fantastique, association facilitant la scolarité d'enfants en situation de handicap et plus particulièrement dyspraxiques, cherche à faciliter la numérisation et l'adaptation d'exercices pour ces enfants. Ce projet a par ailleurs également pour objectif d'être réutilisé par d'autres associations s'engageant pour la même cause.

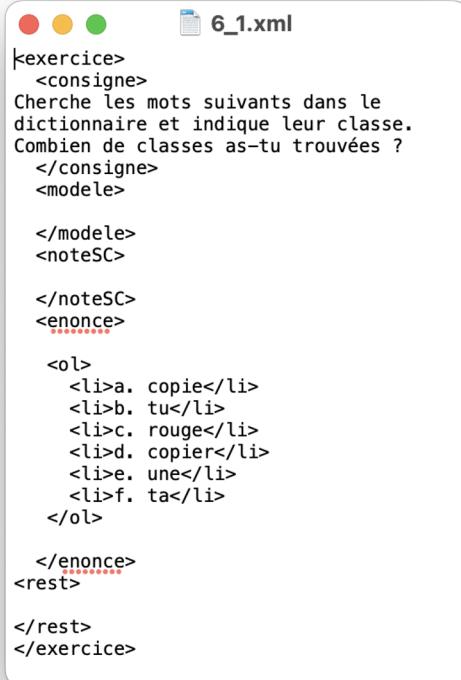
La seule ressource disponible aujourd'hui est constituée de pdfs de manuels scolaires fournis par les éditeurs. Actuellement, les adaptations d'exercices se font manuellement à l'aide d'un outil informatique pour faciliter cette adaptation. L'objectif du projet est donc d'automatiser cette adaptation.



Architecture

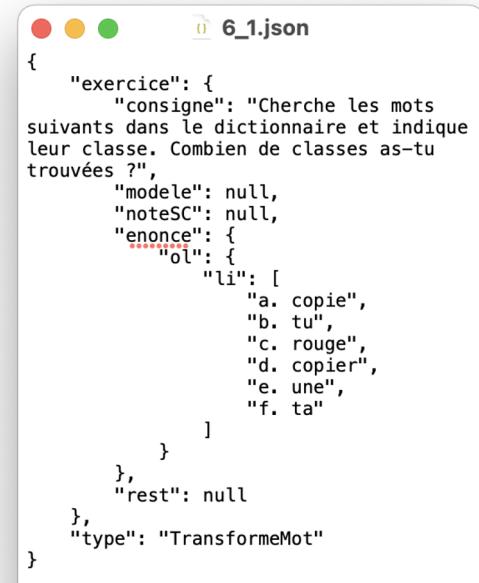
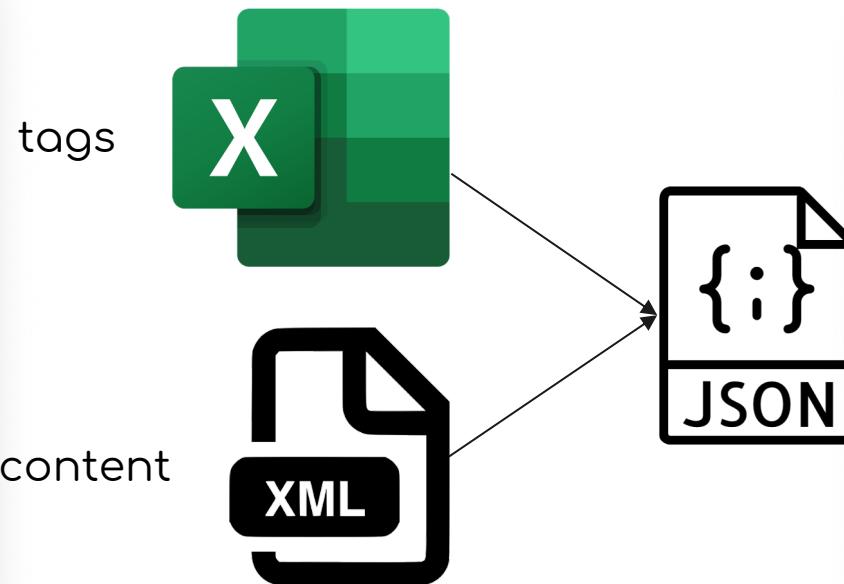


Traitement des données



6_1.xml

```
<exercice>
  <consigne>
    Cherche les mots suivants dans le
    dictionnaire et indique leur classe.
    Combien de classes as-tu trouvées ?
  </consigne>
  <modele>
  </modele>
  <noteSC>
  </noteSC>
  <enonce>
  .....
  <ol>
    <li>a. copie</li>
    <li>b. tu</li>
    <li>c. rouge</li>
    <li>d. copier</li>
    <li>e. une</li>
    <li>f. ta</li>
  </ol>
  <enonce>
  .....
  </enonce>
<rest>
</exercice>
```



6_1.json

```
{
  "exercice": {
    "consigne": "Cherche les mots
    suivants dans le dictionnaire et indique
    leur classe. Combien de classes as-tu
    trouvées ?",
    "modele": null,
    "noteSC": null,
    "enonce": {
      "ol": [
        "li": [
          "a. copie",
          "b. tu",
          "c. rouge",
          "d. copier",
          "e. une",
          "f. ta"
        ]
      },
      "rest": null
    },
    "type": "TransformeMot"
  }
}
```

Diagramme de Classe



Adaptation des exercices

1081 exercices
52 catégories

7 * Termine les phrases avec le complément qui convient.

à une bonne croissance ♦ à l'étage ♦

des plats épicés ♦ à la nuit

a. Le jour succède

b. L'exercice physique contribue

c. Au restaurant, on a goûté

d. Il faut prendre l'ascenseur pour accéder ...

2 * Recopie uniquement les verbes conjugués à l'imparfait.

il cassait ♦ je cherche ♦ nous mangions ♦

tu imaginais ♦ nous finissions ♦ elle tremble ♦

vous passez ♦ elles sonnent ♦ nous

déménageons ♦ elles refroidissaient ♦

vous ponciez ♦ nous rinçons

EDITO

Termine les phrases avec le complément qui convient.

[à une bonne croissance] [à l'étage] [des plats épicés] ou [à la nuit]

a. Le jour succède à une bonne croissance.

[à une bonne croissance] [à l'étage]
[des plats épicés] [à la nuit]

OUPS

EDITO

Colorie uniquement les verbes conjugués à l'imparfait.

[il cassait] ♦ [je cherche] ♦ [nous mangions] ♦ [tu imaginais] ♦ [nous finissions] ♦ [elle tremble] ♦ [vous passez] ♦ [elles sonnent] ♦ [nous déménageons] ♦ [elles refroidissaient] ♦ [vous ponciez] ♦ [nous rinçons]

OUPS

Adaptation - Catégories d'exercices

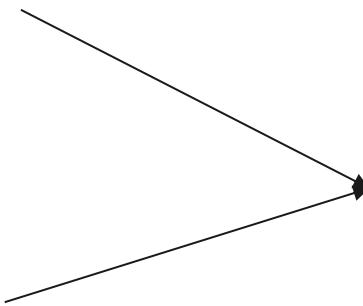
93% des exercices

Select (244)	Fill (556)	Choose (184)	Swap (14)	Show (12)
CocheMots (82)	RC (119)	CM (123)	GroupeEchange (14)	Texte (12)
CocheGroupeMots (18)	EditPhrase (93)	VraiFaux (15)		
CochePhrase (53)	ExpressionEcrite (88)	ClasseCM (8)		
CochelIntrus (11)	TransformePhrase (85)	Associe (38)		
CachelIntrus (29)	TransformeMot (82)			
Classe (51)	RCHandleImage (42)			
	RCCadre (30)			
	RCDouble (17)			

Adaptation - Architecture



Chargement du template



Récupération des données dans le json

Adaptation spécifique à chaque type d'exercice

→ Ecriture dans le fichier html



Select

Trois grandes étapes dans l'adaptation :

- Délimitation des entités sélectables
- Détermination du nombre de couleurs nécessaires
- Adaptation de la consigne

Select

1) Délimitation des entités sélectables

Select : Délimitation des entités sélectables

Trois types d'entités cochables :

① Les phrases

entités séparées par “.”, “!” ou “?”.

② Les groupes de mots

entités séparées par des symboles.

③ Les mots + ponctuation

Select : Délimitation des entités sélectables

Coche Phrases : toujours séparé en phrases.

Colorie en jaune les phrases quand le mot en gras est un nom.

- a. **La** fillette joue dans le jardin.
- b. Il pose la joue sur l'oreiller et s'endort.
- c. Le facteur livre un colis.

Select : Délimitation des entités sélectables

Coche Groupe Mots / Coche intrus / Cache intrus :

Si il y a des symboles dans l'énoncé : séparé selon les symboles.

3 * Relève les groupes de mots qui indiquent le passé.

l'année dernière ◆ dans trente secondes ◆
l'année prochaine ◆ actuellement ◆
dans deux jours ◆ hier ◆ l'été précédent ◆
mercredi dernier

Colorie en jaune les groupes de mots qui indiquent le passé.

l'année dernière dans trente secondes l'année prochaine

actuellement dans deux jours hier l'été précédent

mercredi dernier

Select : Délimitation des entités sélectables

Coche Groupe Mots / Coche intrus / Cache intrus :

Si il y a des symboles dans l'énoncé : séparé selon les symboles.

Sinon : séparé par phrases.

1 * Recopie uniquement les groupes nominaux au pluriel.

- a. le méchant garnement
- b. de hauts immeubles
- c. une souris verte
- d. des pies bavardes
- e. des pluies torrentielles
- f. un vieux fusil

Colorie en jaune uniquement les groupes nominaux au pluriel.

- a. le méchant garnement
- b. de hauts immeubles
- c. une souris verte

Select : Délimitation des entités sélectables

Coche Mots :

Si il y a des symboles dans l'énoncé : séparé selon les symboles.

2 * Dans cette liste, relève les verbes à l'infinitif.

fin ♦ savoir ♦ boisson ♦ attendre ♦ choisir ♦ jeu
♦ rencontre ♦ chanter ♦ sel

Dans cette liste, colorie en jaune les verbes à l'infinitif.

fin savoir boisson attendre choisir jeu rencontre chanter sel

Select : Délimitation des entités sélectables

Coche Mots :

Si il y a des symboles dans l'énoncé : séparé selon les symboles.

Sinon : séparé en mots + ponctuation.

2 * Recopie les noms et les verbes de ce texte.

En 1973, les pays producteurs de pétrole décident brusquement de limiter leur production. Le prix du pétrole s'envole : c'est le premier choc pétrolier. Le gouvernement français lance la construction de 7 centrales nucléaires pour remplacer les centrales électriques classiques.

Colorie en jaune les noms et les verbes de ce texte.

En 1973, les **pays** producteurs de pétrole **décident** brusquement de limiter leur production.

Le prix du **pétrole** s'envole : c'est le premier choc pétrolier.

Select : Délimitation des entités sélectables

Classe :

Si il y a des symboles dans l'énoncé : séparé selon les symboles.

2 * Classe les mots en trois groupes :

nom, verbe, adjectif.

verrou ♦ baigner ♦ joli ♦ chaleur ♦ grosse ♦
surveiller ♦ degré ♦ librairie ♦ repas ♦ parler

Colorie de la bonne couleur les mots : nom, verbe, adjectif.

verrou baigner joli chaleur grosse surveiller degré librairie

repas parler

Select : Délimitation des entités sélectables

Classe :

Si il y a des symboles dans l'énoncé : séparé selon les symboles.

Si il y a "phrase" dans la consigne : séparé en phrases.

2 * Pour chaque phrase, écris si l'action se situe dans le passé, le présent ou le futur.

- a. À ta naissance, tu avais 300 os.
- b. Adulte, tu n'auras plus que 206 os car beaucoup vont se souder.
- c. Nous avons perdu nos dents de lait à l'âge de 6 ans.

Pour chaque phrase, écris si l'action se situe dans le passé, le présent ou le futur.

a. À ta naissance, tu avais 300 os.

b. Adulte, tu n'auras plus que 206 os car beaucoup vont se souder.

Select : Délimitation des entités sélectables

Classe :

Si il y a des symboles dans l'énoncé : séparé selon les symboles.

Si il y a "phrase" dans la consigne : séparé en phrases.

Sinon : séparé en mots + ponctuation.

**3 ** Relève dans ce texte les verbes conjugués et les verbes à l'infinitif.
Classe-les dans le tableau.**

verbes conjugués	verbes à l'infinitif

La tortue leva ses yeux noirs et enfoncés sur le petit garçon. [...] Personne ne bougeait. Puis, avec une grande dignité, l'énorme bête se retourna et se dirigea vers le bord de l'eau en se dandinant, sans se presser. Elle traversa posément la plage de sable et sa grosse carapace se balançait doucement. La foule regardait en silence. La tortue entra dans l'eau. Elle continua d'avancer.

Relève dans ce texte les verbes conjugués et les verbes à l'infinitif. Classe-les

: **verbes conjugués** **verbes à l'infinitif**.

La tortue leva ses yeux noirs et enfoncés sur le petit garçon .

[...] Personne ne bougeait .

Select : Délimitation des entités sélectables

Limites :

- Possibilité de présence de symboles qui ne sont pas des séparateurs
- Les règles pour répartir les exercices ne sont pas parfaites

Select

2) Détermination du nombre de couleurs

Select : Détermination du nombre de couleurs



Cas où les couleurs sont dans la consigne

“Recopie les phrases. Souligne en bleu le verbe avoir conjugué au présent et en vert le mot invariable à.”

→ Ces couleurs sont celles affichées

Liste des couleurs possibles : bleu, vert, noir et rouge (couleurs des bics)

Select : Détermination du nombre de couleurs

“Souligne les deux GN dans chaque phrase. Encadre le nom principal de chaque groupe.”

Select : Détermination du nombre de couleurs

“Souligne les deux GN dans chaque phrase. Encadre le nom principal de chaque groupe.”

V

V

Technologie utilisée : Modèle de NLP "gilf/french-postag-model", Hugging face
Tokenisation et POS tagging



HUGGING FACE

Select : Détermination du nombre de couleurs

“Souligne les deux GN dans chaque phrase. Encadre le nom principal de chaque groupe.”

V

NC

NC

V

NC

NC

Select : Détermination du nombre de couleurs

“Souligne les deux GN dans chaque phrase. Encadre le nom principal de chaque groupe.”

V

NC

NC

V

NC

NC

min (V, NC)

Select : Détermination du nombre de couleurs

“Souligne les deux GN dans chaque phrase. Encadre le nom principal de chaque groupe.”

V

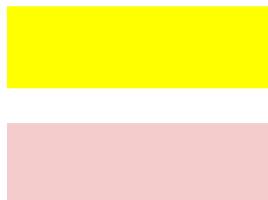
NC

NC

V

NC

NC



Note : les auxiliaires
ne sont pas
comptés

Modèle de Hugging face : les auxiliaires sont tagguées comme des verbes



Utilisation de **spaCy** pour détecter les auxiliaires et les retirer

Select : Détermination du nombre de couleurs



Affinage

“Recopie uniquement les déterminants qui marquent le pluriel.”

V

NC

V

NC

$\min(V, NC) = 2$
ERREUR

Select : Détermination du nombre de couleurs



Affinage

“Recopie uniquement les déterminants qui marquent le pluriel.”

V

NC

V

NC

$$\min(V, NC) = 2$$

ERREUR

Select : Détermination du nombre de couleurs



Affinage

“Recopie uniquement les déterminants qui marquent le pluriel.”

V

NC

V

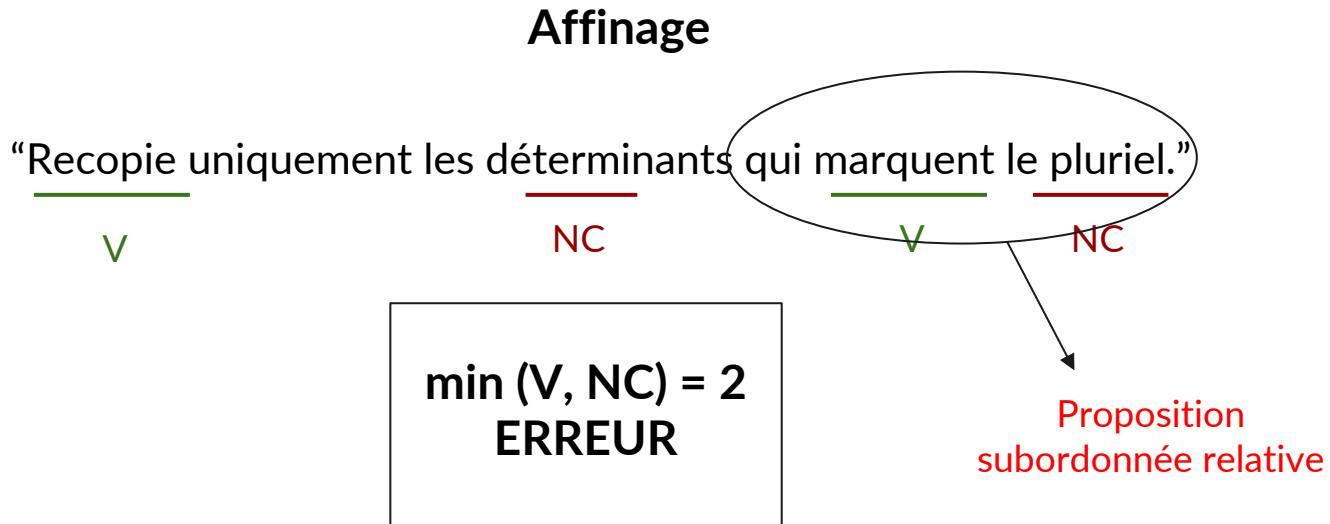
NC

$$\min(V, NC) = 2$$

ERREUR

Proposition
subordonnée relative

Select : Détermination du nombre de couleurs



→ Pas de prise en compte du premier verbe situé après un pronom relatif “qui”, “qu”, “dans lequel”, “dans lesquels”, ...

Select : Détermination du nombre de couleurs

Bilan :

- Couleurs présentes dans la consigne : pas de difficulté
- Couleurs non présentes dans la consigne : POS tagging

Limites :

- Erreurs de POS tagging
- Tous les verbes “inutiles” ne sont pas dans des propositions subordonnées relatives



Select

3) Adaptation de la consigne

Select : Adaptation de la consigne

Radicaux des verbes : “**encadr**”,
“**entour**”, “**soulign**”, “**recop**”, “**relev**”.

Pour le type Classe: “**class**”

Si un mot possède ce radical :
on le remplace par “**Colorie en**” + la
couleur

Si un mot possède ce radical :
on le remplace par “**Colorie de la bonne
couleur**”

Select : Adaptation de la consigne

Entoure le pronom personnel et
souligne le nom qu'il remplace.

→ Colorie en jaune le pronom personnel et
colorie en rose le nom qu'il remplace.

Souligne les compléments.

→ Colorie en jaune les compléments.

Souligne en bleu le verbe avoir et en
vert le mot invariable à.

→ Colorie en bleu le verbe avoir et en vert
le mot invariable à.

Classe les noms selon leur nombre.

→ Colorie de la bonne couleur les noms
selon leur nombre.

Select : Adaptation de la consigne

Cas particulier des types “Intrus”

Une seule phrase dans la consigne

14 * Recopie chaque liste sans l'intrus.

- a. partout ♦ grandes ♦ aussi ♦ demain
- b. tôt ♦ longtemps ♦ dormir ♦ assez
- c. espadrille ♦ pour ♦ de ♦ par
- d. car ♦ donc ♦ mais ♦ suivant

Dans chaque liste, cache l'intrus

- a.

Plusieurs phrases dans la consigne

2 * Chaque liste contient une forme

du verbe *faire* ou *aller* qui n'est pas au futur. Recopie les listes sans l'intrus.

- a. je ferai ♦ il ira ♦ nous ferons ♦ ils font
- b. tu vas ♦ vous irez ♦ elle fera ♦ j'irai
- c. vous ferez ♦ nous irons ♦ elle irait ♦ tu feras
- d. nous ferions ♦ elle ira ♦ je ferai ♦ tu iras

Chaque liste contient une forme du verbe faire ou aller qui n'est pas au futur. Dans chaque liste, cache l'intrus

- a.

Select : Adaptation de la consigne

Limites :

- Possibilité de présence de verbes non répertoriés dans l'algo
- Les Intrus : la phrase générique peut faire perdre des informations si celles-ci ne sont pas données dans une phrase à part.

Select : Adaptation de la consigne

Cas particulier du type “Classe”

Couleurs dans la consigne

Colorie de la bonne couleur les mots : nom, verbe, adjectif .

111_9

Select : Adaptation de la consigne

1

Cas particulier du type “Classe”

“Classe les mots dans le tableau selon qu’ils contiennent le son [g] ou le son [ʒ].”

Fonction qui trouve une liste de choix dans la
consigne → [“le son [g]”, “le son [ʒ]”]

Select : Adaptation de la consigne

1

Cas particulier du type “Classe”

“Classe les mots dans le tableau selon qu’ils contiennent le son [g] ou le son [ʒ].”

Fonction qui trouve une liste de choix dans la
consigne → [“le son [g]”, “le son [ʒ]”]

Colorie de la bonne couleur les mots selon qu’ils contiennent le son [g] ou le son [ʒ].

Select : Adaptation de la consigne

2

Cas particulier du type “Classe”

“Classe les mots selon la prononciation du e. [e] comme bébé [ɛ] comme terre”

P111Ex6

Select : Adaptation de la consigne

2

Cas particulier du type “Classe”

“Classe les mots selon la prononciation du e. [e] comme bébé [ɛ] comme terre”

P111Ex6

Ajout d'une
balise dans le
xml

```
cartable-fantastique-fall-2021-p1-data > data > exs > 111_6.xml
1  <exercice>
2  |  <consigne>
3  |  Classe les mots selon la prononciation du e.
4  |  </consigne>
5  |  <categories>
6  |  [e] comme bébé. [ɛ] comme terre.
7  |  </categories>
8  |  <modele>
9
10 </modele>
```

Select : Adaptation de la consigne

2

Cas particulier du type “Classe”

“Classe les mots selon la prononciation du e. [e] comme bébé [ɛ] comme terre”

P111Ex6

Modification de la consigne en y ajoutant “:” suivi des catégories

Colorie de la bonne couleur les mots selon la prononciation du e :

[e] comme bébé

[ɛ] comme terre



Select : Statistiques

STATS

Coche Phrases : 53 exercices	→	92%	Nombre d'exercices bien adaptés
Coche Groupe Mots : 18	→	75%	74,2%
Coche Mots : 82 exercices	→	61,8%	
Coche Intrus : 11	→	81,8%	155 exercices bien adaptés sur 244
Cache Intrus : 29 exercices	→	71,4%	
Classe : 51 exercices	→	72,2%	

Swap (GroupeEchange)

- Détermination des entités qui seront les étiquettes à déplacer : toujours séparation des entités selon des symboles (♦)
- 14 exercices → 6 bien adaptés, 3 cas impossibles, 3 mal extraits

Fill : tous les remplissages au clavier

- Détermination des endroits à remplir au clavier
- Adaptation de la consigne

Fill : Détermination des endroits à remplir

Pré-Traitement

Texte sous
forme de
sentences (blocs
insécables) :
`list(str)`

`text_to_html`

Post-Traitement

Sous la forme
de html
Remplacement
par dictionnaire
de valeurs à
remplacer

Dictionnaire de valeurs
à remplacer

Fill : Détermination des endroits à remplir



RemplirClavier : Repérage des différents symboles à remplacer par une box éditable (... , , _) et d'autres symboles à remplacer/supprimer (♦, etc)

TransformePhrase : Détection des phrases

EditPhrase : Récupération de la phrase à modifier (en enlevant les bouts parasites comme a., b.)

Complète chacune de ces phrases avec un adjectif.

a. Ce nouveau canapé est très .

Reconstitue la phrase. N'oublie pas la majuscule ni la ponctuation.

Écris ces phrases au futur.

b. Le TGV est le train le plus de France.

Hugo ♦ nous avons invité ♦ pour le goûter

a. Tu veux faire une sortie en forêt.

c. Ce serpent est très .

→

→ Tu veux faire une sortie en forêt.

Fill : Détermination des endroits à remplir

RCCadre : Détection du ou des verbes, à partir des délimiteurs, très souvent () dans ce manuel

ExpressionEcrite : Affichage simple

RCImage : Pas traitable

RCDouble : Récupération des mots avec un style différent (gras, italique, souligné, etc)

Recopie les phrases avec le verbe conjugué au présent.

venir

a. Il me voir.

10 **À l'oral** Avec un camarade, présentez la signification de chaque panneau au reste de la classe.



18 *** Réécris ce texte en mettant les verbes en gras au futur.

Enfin il **peut** faire ce qui lui plaît en sortant de l'école. Ce qu'il préfère, c'est jouer au football. Mais il **est** tout seul. Les autres ne **viennent** au stade que lorsqu'ils **ont** terminé leurs devoirs. Paul **s'allonge** dans l'herbe et **regarde** le ciel. Paul **s'ennuie**.

Hans Joachim Schädlich,
Le Coupeur de mots,
trad. J. Etoré et B. Lortholary,
© Flammarion, coll. « Castor Poche »



Fill : Détermination des endroits à remplir

TransformeMot : Distinction de différents sous-types d'exercices :

- Liste de verbes, avec une indication de temps et personnes de conjugaison : Détection du séparateurs entre les verbes, reformatage pour l'affichage
- Liste de mots séparés en général par des ♦ : Reformatage du texte pour l'affichage, récupération des éventuels endroits à remplir ou non
- Cas plus classiques de listes a. b. c. : Reformatage de l'affichage

9 * **Conjugue les verbes à l'imparfait à la personne indiquée.**

- a. *aller; dire* → 1^{re} personne du singulier
 b. *faire; voir* → 2^e personne du singulier
 c. *pouvoir; vouloir* → 3^e personne du singulier
 d. *venir; prendre* → 1^{re} personne du pluriel

3 * **Indique à quelle classe grammaticale appartiennent ces mots.**

- ouverture ♦ énorme ♦ un ♦ dormir ♦
 téléviseur ♦ être ♦ écouter ♦ chameau ♦
 somptueux ♦ magique ♦ chat ♦ magnifique ♦
 travailler ♦ manger ♦ voiture

5 ** **Trouve le verbe qui correspond à chacun de ces noms.**

- a. la baignade
 b. le rugissement
 c. la vente
 d. le chant

Fill : Adaptation de la consigne

Pour RCCadre : Il n'y a plus de verbe entre parenthèses

Conjugue au présent les verbes.

6 ★ Conjugue au présent les verbes entre parenthèses.

- a. Nous (*apprendre*) les mathématiques.
- b. tu (*comprendre*) ce que je dis ?
- c. Je (*reprendre*) du gâteau s'il en reste.
- d. Tu (*surprendre*) le chat dans sa cachette.

apprendre

a. Nous les mathématiques.

Fill : Statistiques

Catégorie	Nb d'exs impossible	bien adaptés	mal adaptés	précision
RC	119	3	115	1 99,1%
EditPhrase	93	19	73	1 98,6%
TransformePhrase	85	15	65	5 92,9%
TransformeMot	82	14	67	1 98,5%
RCCadre	30	4	26	100,0%
<i>ExpressionEcrite</i>	88	88		-%
<i>RCImage</i>	42	42		-%
<i>RCDouble</i>	17	17		-%
Total	556	202	346	8 97,74%

Choose : Détails de la conversion

- Repérage des choix
- Préparer à l'affichage
- Préparation des choix pour la conversion
- Adaptation de la consigne
- Classe Associe



Choose : Détails de la conversion

1) Repérage des choix

Choose : Recherche des choix (consigne)

“Écris ces phrases en les terminant par un point (.), un point d'interrogation (?) ou un point d'exclamation (!).”

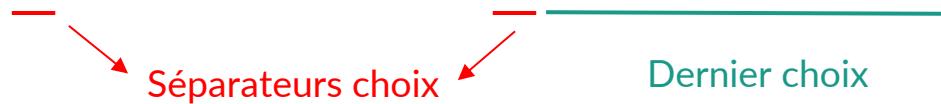


HYPOTHÈSES:

- Les choix sont séparés par des séparateurs généralisables à toutes les consignes
- Les choix sont tous localisés dans une seule phrase

Choose : Recherche des choix (consigne)

“Écris ces phrases en les terminant par un point (.), un point d'interrogation (?) ou un point d'exclamation (!).”



Choose : Recherche des choix (consigne)

“Écris ces phrases en les terminant par un point (.), un point d'interrogation (?) ou un point d'exclamation (!).”



Choose : Recherche des choix (consigne)

“Écris ces phrases en les terminant par un point (.), un point d'interrogation (?) ou un point d'exclamation (!).”

Premier choix

Choix intermédiaires

Séparateurs choix

Dernier choix

HYPOTHÈSE:

- Indicateur du lieu du premier choix : dans le même nombre de mots que le dernier

Donc ici, dans “par un point (.)” (4 mots)

Choose : Recherche des choix (consigne)

["par", "un", "point", "(.)"]

Candidat Premier choix

["un", "point", "d'exclamation", "(!)"]

Dernier choix

Heuristique de comparaison:

{

- Comparer tous les mots sauf le dernier
- score = 0
- si mot dans dernier choix, score = score
- sinon score = score + Levenshtein(mot, même index dernier choix)
- si score < seuil * nombres caractères dernier choix, candidat valide

Choose : Recherche des choix (sous-consigne)

“à une bonne croissance ◆ à l'étage ◆ des plats épicés ◆ à la nuit”

Séparateurs choix

Par défaut, l'espace est un séparateur

[“à”, “une”, “bonne”, “croissance”, “”, “”, “à”, “l'étage”, “”, “”, “des”, “plats”, “épicés”, “”, “”, “à”, “la”, “nuit”]

Possibilité de repérer les séparateurs

Choose : Recherche des choix (énoncé)

“a. Tu (auras/aura) besoin d'aide.”

— — —

Symboles communs aux phrases

Remplacement par un symbole unique “§”,

“a. Tu §§auras§aura§§ besoin d'aide.”

Traitement différent selon qu'il y a un symbole de remplissage (ex: "...") dans les phrases ou non

HYPOTHÈSE:

- Si les choix ne se trouvent pas dans une des phrases de l'énoncé, on peut la traiter comme une recherche dans la sous-consigne

Choose : Recherche des choix

Différents cas où les choix ne sont pas trouvés:

- Le lièvre/Le lièvre et la tortue font la course. → Le lièvre et la tortue font la course. (21_7)
- Sous-consigne du (31_7)

7

** Complète le texte avec des articles.



Utilise 1 fois une, 2 fois le, 4 fois la,
2 fois l', 2 fois les et 1 fois des.

- Complète les mots avec une consonne simple ou une consonne double. (111_8)
→ les consonnes se trouvent dans l'énoncé mais l'algorithme repère des choix dans la consigne
- Classe les groupes nominaux dans le tableau. article défini article indéfini (38_6)

Choose : Recherches des choix

Bilan :

- Très bonnes performances sur le manuel de test
- Possibilité de customizer plusieurs paramètres de la recherche
- Pleins de pistes à explorer pour améliorer les résultats

Limites :

- Certaines hypothèses ne sont pas généralisables et contredites par la nouvelle sortie de l'extraction → adaptation en conséquence du code
- Certains patterns ne peuvent juste pas être reconnus → utilisation de la nouvelle sortie de l'extraction pour y pallier



Choose : Détails de la conversion

2) Préparer à l'affichage

Choose : Préparer à l'affichage

- Rechercher les différentes lignes à afficher
- Rechercher les différents symboles à remplir (ex: "...") et dans le cas contraire déterminer où les placer :

- 
- par défaut : "line → ..."
 - début de ligne, mot différent : "... line"
 - début de ligne : "...line"
 - fin de ligne : "line..."
 - fin de ligne, mot différent : "line ..."



Choose : Préparation à l'affichage

Bilan :

- 100 % de bonne préparation si les choix sont bien repérés



Choose : Détails de la conversion

3) Préparation des choix pour la conversion

Choose : Nettoyage des choix

- Retirer les tirets “-” en début et fin (ex: “im-”)
- Repérer d'éventuels symboles comprenant les symboles (ex: “(?)”)
- Retirer en amont les prépositions et les articles contractés (ex: “au”, “de”, “d”, etc...) si les choix se situaient dans la consigne

Choose : Mettre en majuscule

- Nécessité de parcourir les lignes de l'énoncé à afficher
- Regex pour savoir quand mettre une majuscule
- 100 % de réussite de l'algorithme

Choose : Préparation des choix pour la conversion

Bilan :

- Bonne performance du nettoyage
- Trouver quand mettre une majuscule est déterministe

Limites :

- Retirer les prépositions et les articles contractés peut poser problème dans certains cas



Choose : Détails de la conversion

4) Adaptation de la consigne

Choose : Adaptation de la consigne

100 % de bonne conversions sur la base de test

- 
- Remplacer les “recopie” par “complète”
 - Remplacer les “écris” et “réécris” par “choisis”
 - Retirer les phrases courtes avec les précédents verbes (ex: Ecris ces phrases.)

Choose : Adaptation de la consigne

Bilan :

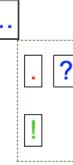
- Adaptation de la consigne : pas de difficulté

Choose : Rendu graphique

Choisis ces phrases en les terminant par un point (.), un point d'interrogation (?) ou un point d'exclamation (!).

. ? ou !

a. Il sonne à la porte et entre ...



>

Rendu graphique de l'exercice 17_9 tagué ChoixMultiples

Choose : Cas de la classe Associe

<enonce>

Thibaud s'envolent Les mouettes jouent Les enfants court L'athlète dort

</enonce>

4 ★ Associe chaque sujet à son verbe.

- | | | |
|--------------|-----------------------|----------------------------------|
| Thibaud | <input type="radio"/> | <input type="radio"/> s'envolent |
| Les mouettes | <input type="radio"/> | <input type="radio"/> jouent |
| Les enfants | <input type="radio"/> | <input type="radio"/> court |
| L'athlète | <input type="radio"/> | <input type="radio"/> dort |

- Impossibilité de convertir avec la pipeline de choose
- Impossibilité de convertir avec l'ancienne version de l'extraction

Choose : Statistiques



STATS

Choix Multiples : 92 exercices	→	97%
Classe CM : 8 exercices	→	75%
VraiFaux : 13 exercices	→	100%
Associe : 37 exercices	→	?%

Nombre d'exercices bien adaptés

96 % (sans Associe)

108 exercices bien adaptés sur 113 (sans Associe)

Adaptation : Statistiques

Catégorie	Nb d'exe impossible	bien adaptés	mal adaptés	précision
Fill	556	202	346	8 97,7%
Select	244	35	155	54 74,2%
Choose	184	71	108	5 95,6%
Swap	14	3	6	5 54,5%
Show	12	11	1	100,0%
Non Traités	71	14	56	0,0%
Total	1081	336	616	128 82,68%

Rendu Graphique



EDITO

Complète chaque phrase avec un adjectif de ton choix.

- a. Au printemps, l'air est généralement .
- b. La nature émerge d'un sommeil hivernal.
- c. Arrête de bouger, reste une minute.

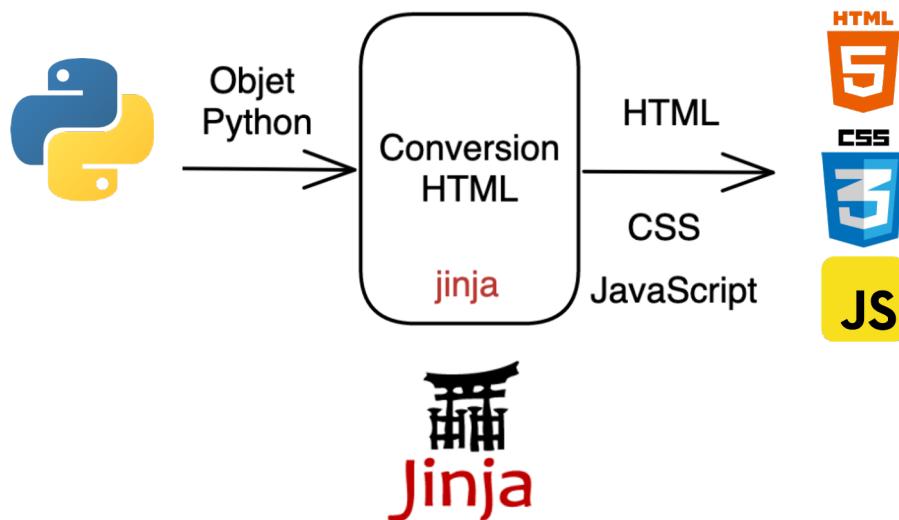


OUPS

Complète chaque phrase avec un adjectif de ton choix.

- a. Au printemps, l'air est généralement .
- b. La nature émerge d'un sommeil hivernal.
- c. Arrête de bouger, reste une minute.

Rendu Graphique : Architecture



- L'adaptation donne un code html grâce à la fonction `text_to_html`
- Ce code est transformé en fichiers html grâce à Jinja qui remplit des templates
- Le javascript donne les couleurs, sépare sur plusieurs pages pour chaque fichier html
- Il y a un fichier js et un fichier css global + éventuellement un par type d'exercice

Rendu Graphique

Objectif => être le plus fidèle possible au site cahiers fantastiques

Features :

- Espacement important
- Lignes de couleurs différentes (alternance bleu/rouge/vert)
- Nombre maximum de lignes par pages (dépendant des exercices)
- Éléments graphiques spécifiques à des types d'exercice (choix, colorie, etc)

Feature : Espacement important

```
▼<body>
  <button id="previous_page" style="display: none;"></button>
  <button id="next_page" style="display: inline;"></button>
  ▼<div id="central_block">
    ▶<div id="guideline">...</div>
    <div id="additional_guideline"></div>
    ▶<div id="exercise_text">...</div>
  </div>
```



```
▼<span id="word2" class="blue">
  <span class="word">printemps,</span>
  <span class="space"> </span>
</span>
```

```
.space {
  white-space: pre-wrap;
}
```

Rendu Graphique

Feature : Lignes de couleurs différentes

- Parcours de chaque mot
- Récupération de la position verticale
- Détection de lignes (à quelques pixels près)
- Attribution d'une couleur par ligne

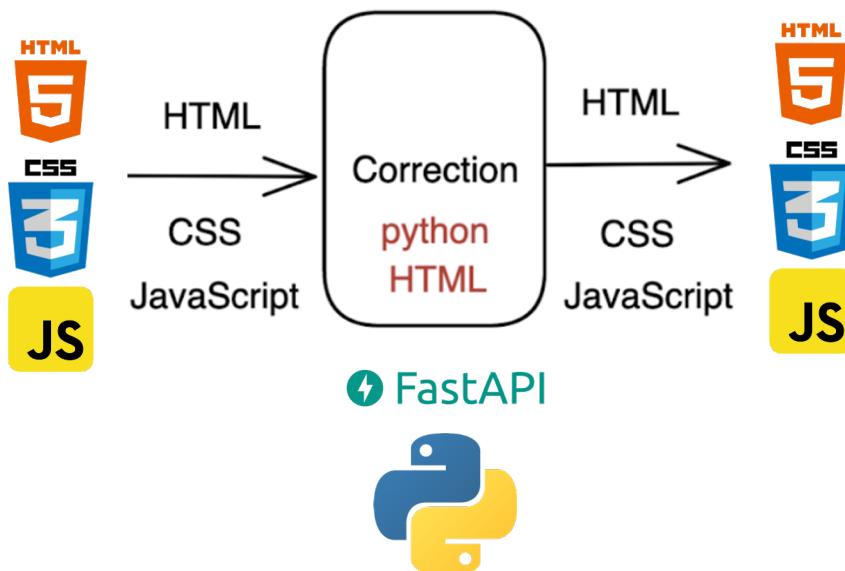
Feature : Gestion des pages

- Séparation de chaque partie insécable en "block"
- Détection du nombre de lignes qu'occupe un block
- Calcul de où peut être placé un block
- Positionnement du block sur la bonne page (avec création de la page)
- Apparition/disparition des boutons pour changer de page

```
<div id="exercise_text">  
  <p id="page0" style="display: block;">  
    <span class="block" id="block0">...</span>  
    <span class="block" id="block1">...</span>  
    <span class="block" id="block2">...</span>  
  </p>  
  <p id="page1" style="display: none;">...</p>  
</div>
```

remplir_clavier/37_10.html#page0

Interface de correction : Architecture



- Un serveur python FastAPI fait tourner le service de correction
- Le serveur modifie directement des fichiers sur la machine hôte
- L'instance de service est locale

Interface de correction : Architecture

Objectif => pouvoir corriger les html en sortie de la pipeline

Features :

- Parcourir les exercices (avec les touches directionnelles)
- Indiquer si l'exercice est bien converti ou non
- Indiquer si l'exercice est mal extrait
- Modifier le tag de l'exercice et relancer la conversion

Choisis ces phrases en les terminant par un point (.), un point d'interrogation (?) ou un point d'exclamation (!).

[.] [?] ou [!]

b. Entrez ou sortez, mais fermez la porte ...

Bonne Conversion Mauvaise Conversion

Nouveau tag: ChoixMultiples Générer avec nouveau tag

Mauvaise Extraction

Aperçu de l'interface de correction d'un exercice

Taggage des exercices : Données

Données :

- Obtention des 1060 exercices avec un tag
- Récupération des xmls bruts, ou des xmls sans balises
- On ne garde qu'un certain nombre de classes, et les autres exercices sont dans une classe "Autres"

		text	labels
0		Cherche les mots suivants dans le dictionnaire...	TransformeMot
1		Classe les mots en trois groupes : nom, verbe,...	Classe
2		Complète avec : le, une, un, des, tu, elles, i...	Autres
3		Écris une phrase en respectant l'ordre des cla...	ExpressionEcrite
4		Recopie les mots suivants, puis entoure les pr...	Autres

Taggage des exercices : Traitement

Utilisation de SimpleTransformers

Variation selon les données utilisées :

(Camembert,
16 classes, 10 epochs)

Deux modèles pré-entraînés utilisés :

- Camembert
- Flaubert



	precision	recall	f1-score	support		precision	recall	f1-score	support
TransformeMot	0.68	0.68	0.68	19	RC	0.77	0.92	0.84	25
RCImage	1.00	0.83	0.91	6	RCDouble	0.00	0.00	0.00	8
TransformePhrase	0.42	0.53	0.47	15	TransformeMot	0.72	0.87	0.79	15
CochePhrase	1.00	1.00	1.00	12	RCImage	1.00	0.75	0.86	8
RC	0.86	0.66	0.75	29	Autres	0.69	0.82	0.75	22
CacheIntrus	1.00	1.00	1.00	7	RCCadre	0.67	1.00	0.80	2
Associe	0.67	1.00	0.80	6	TransformePhrase	0.83	0.53	0.65	19
Classe	0.57	0.73	0.64	11	CochePhrase	0.90	1.00	0.95	9
EditPhrase	0.65	0.71	0.68	21	CacheIntrus	0.60	0.75	0.67	4
RCDouble	0.00	0.00	0.00	3	Associe	0.86	0.86	0.86	7
CocheGroupeMots	1.00	0.67	0.80	3	ExpressionEcrite	0.83	0.83	0.83	18
ExpressionEcrite	0.88	0.58	0.70	12	CM	0.92	0.76	0.83	29
RCCadre	0.33	1.00	0.50	1	Classe	0.80	0.80	0.80	10
Autres	0.61	0.63	0.62	30	CocheGroupeMots	1.00	0.43	0.60	7
CM	0.83	0.83	0.83	23	EditPhrase	0.58	0.88	0.70	16
CocheMots	0.95	1.00	0.97	18	CocheMots	0.80	0.94	0.86	17
accuracy			0.74	216	accuracy			0.78	216
macro avg	0.72	0.74	0.71	216	macro avg	0.75	0.76	0.74	216
weighted avg	0.75	0.74	0.74	216	weighted avg	0.77	0.78	0.76	216

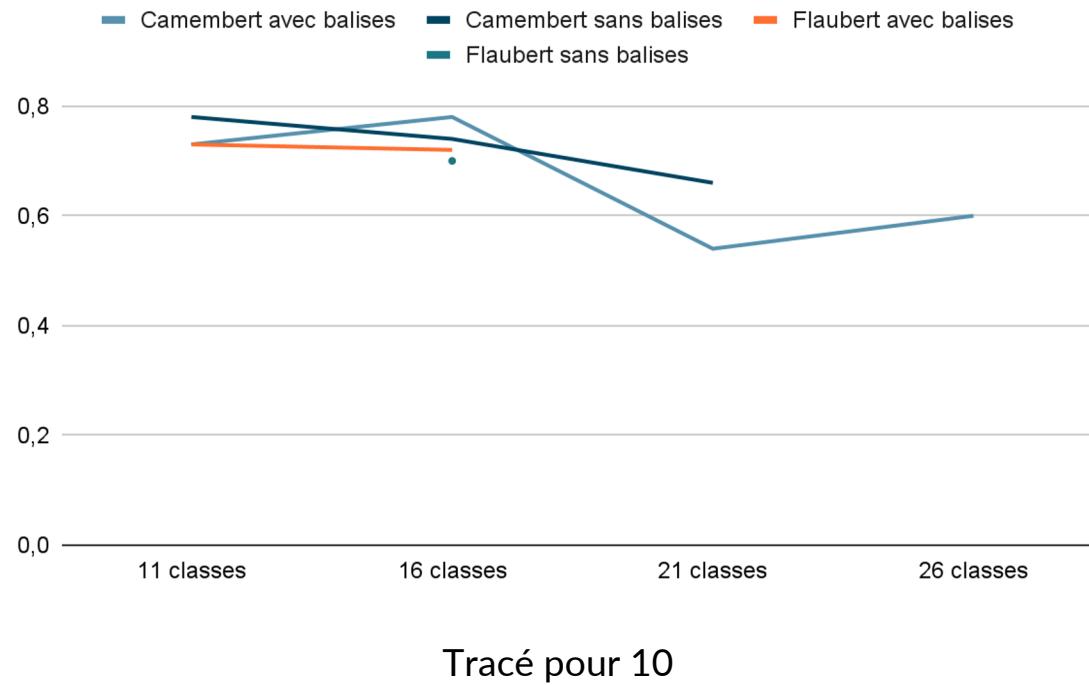
sans balises

avec balises

Taggage des exercices : Traitement



Accuracy



Nombre de classes optimal ?

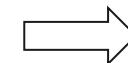
Avec balises vs sans balises :
~ accuracy similaires

Camembert vs Flaubert :
~ accuracy similaires

Taggage des exercices : Traitement

Taggage des catégories d'exercices (Fill, Swap, Show, Choose, Select)

	precision	recall	f1-score	support
swap	0.50	0.50	0.50	4
select	0.88	0.94	0.91	63
show	1.00	0.25	0.40	4
autre	0.46	0.25	0.32	24
choose	0.78	0.75	0.77	48
fill	0.85	0.91	0.88	181
accuracy			0.83	324
macro avg	0.75	0.60	0.63	324
weighted avg	0.82	0.83	0.82	324



plus les catégories sont fournies, plus la précision est grande

Taggage des exercices : Traitement

Taggage des exercices au sein de Select

	precision	recall	f1-score	support
CacheIntrus	0.90	1.00	0.95	9
CocheMots	0.81	0.72	0.76	29
CocheGroupeMots	0.50	0.75	0.60	4
CochePhrase	0.82	0.93	0.87	15
Classe	0.83	0.71	0.77	14
CocheIntrus	0.67	0.67	0.67	3
accuracy			0.80	74
macro avg	0.76	0.80	0.77	74
weighted avg	0.80	0.80	0.80	74

Précision du modèle très limité par le nombre d'exercices ...

Taggage des exercices : Traitement

Variation selon
les données
utilisées :

(Avec Balises,
16 classes, 10
epochs)

	precision	recall	f1-score	support		precision	recall	f1-score	support
TransformeMot	0.80	0.76	0.78	21	RC	0.77	0.92	0.84	25
RCImage	0.67	0.80	0.73	10	RCDouble	0.00	0.00	0.00	8
TransformePhrase	0.50	0.50	0.50	14	TransformeMot	0.72	0.87	0.79	15
CochePhrase	0.88	0.78	0.82	9	RCImage	1.00	0.75	0.86	8
RC	0.88	0.73	0.80	30	Autres	0.69	0.82	0.75	22
CacheIntrus	1.00	1.00	1.00	4	RCCadre	0.67	1.00	0.80	2
Associe	0.50	0.60	0.55	5	TransformePhrase	0.83	0.53	0.65	19
Classe	0.71	0.83	0.77	6	CochePhrase	0.90	1.00	0.95	9
EditPhrase	0.57	0.63	0.60	19	CacheIntrus	0.60	0.75	0.67	4
RCDouble	0.25	0.50	0.33	2	Associe	0.86	0.86	0.86	7
CocheGroupeMots	1.00	1.00	1.00	3	ExpressionEcrite	0.83	0.83	0.83	18
ExpressionEcrite	0.76	0.76	0.76	17	CM	0.92	0.76	0.83	29
RCCadre	1.00	1.00	1.00	7	Classe	0.80	0.80	0.80	10
Autres	0.64	0.52	0.57	27	CocheGroupeMots	1.00	0.43	0.60	7
CM	0.67	0.77	0.71	26	EditPhrase	0.58	0.88	0.70	16
CocheMots	0.88	0.88	0.88	16	CocheMots	0.80	0.94	0.86	17
accuracy			0.72	216	accuracy			0.78	216
macro avg	0.73	0.75	0.74	216	macro avg	0.75	0.76	0.74	216
weighted avg	0.73	0.72	0.72	216	weighted avg	0.77	0.78	0.76	216

flaubert

camembert

Taggage des exercices : Traitement

Flaubert vs Camembert :
~ accuracy similaires

Recall de “Autres” ? (pour 10 epochs)
=> pas concluant

Flaubert		
11 classes	avec balises	0.57
16 classes	avec balises	0.52
	sans balises	0.47

Camembert		
11 classes	avec balises	0.70
	sans balises	0.72
16 classes	avec balises	0.82
	sans balises	0.63
21 classes	avec balises	0.06
	sans balises	0.11
26 classes	avec balises	0.00

Taggage des exercices : Bilan

Limites :

- Faibles nombre de données : 1060 exercices répartis en train set et eval set pour plusieurs dizaines de catégories
- Faibles nombres d'époques d'entraînement : 10, limitations matérielles

Explorations possibles :

- Quel choix de données ? xmls bruts ?
 - Camembert ou Flaubert ?
 - Combien de classes ? Utilisation de grosses classes ?
- => Autres algorithmes : Guided LDA ?

Précision en premier jet : 78%

Organisation du rendu

The image shows two side-by-side screenshots of GitLab project pages. Both projects are named "Cartable Fantastique Fall 2021 p1".

Project Overview:

- Cartable Fantastique Fall 2021 p1:** Project ID: 951. Contains 359 commits, 3 branches, 0 tags, 54.4 MB files, and 56.2 MB storage.
- Cartable Fantastique Fall 2021 p1 DATA:** Project ID: 966. Contains 125 commits, 1 branch, 0 tags, 5 MB files, and 5 MB storage.

Commit History:

- Cartable Fantastique Fall 2021 p1:** Includes a merge commit from 'choose' into 'master' by All Charara 1 day ago, and several other commits from vscode, fantastic, jinja, resources, .gitignore, .gitlab-ci.yml, and .pylintrc.
- Cartable Fantastique Fall 2021 p1 DATA:** Includes a commit generated all expression ercrite by Benjamin Poux 10 hours ago, and several other commits from data, output, pitch-demo, sorted_output, .gitignore, README.md, and others.

Sidebar:

- Common sidebar items: Project overview, Details, Activity, Releases, Repository, Issues (0), Merge requests (0), CI/CD, Security & Compliance, Operations, Packages & Registries, Analytics, Wiki, Snippets, Members, Collapse sidebar.
- Specific to the DATA project: Set up CI/CD.

Conclusion

- Un rendu graphique satisfaisant
- Une interface de correction facilitant l'adaptation
- Un grand nombre d'exercices bien adaptés : 83%
- Une possibilité de tagger automatiquement les exercices avec plus de données et de ressources matérielles

Merci pour votre attention Questions ?

Vos contacts :

Charlotte SASSON, Etudiante Paris Digital Lab, CentraleSupélec, charlotte.sasson@student-cs.fr, 06 87 20 08 02

Ali CHARARA, Etudiant Paris Digital Lab, CentraleSupélec, ali.charara@student-cs.fr, 06 67 42 04 11

Benjamin POUX, Etudiant Paris Digital Lab, CentraleSupélec, b.pouxboiret@gmail.com, 07 68 42 58 97

Renaud MONNET, Référent Digital Lab, renaud.monnet@centralesupelec.fr, 06 12 80 60 40

Théo RUBENACH, Coach Technique Paris Digital Lab, theo.rubenach@illuin.tech, 06 04 45 10 14