

Test Technique Data Scientist - GED

Durée : 5h

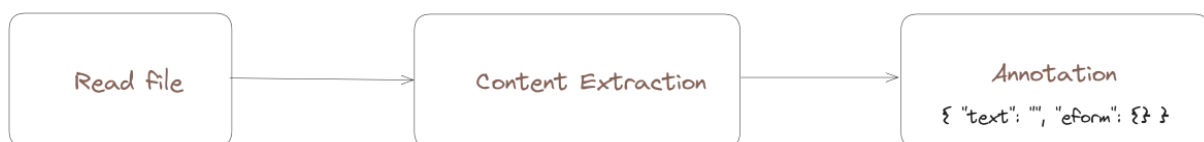
Contexte :

Vous recevez un ensemble de documents administratifs typiques pour une collectivité (actes de naissance, passeports, certificats, ...). Votre tâche consiste à :

1. Extraire des données structurées à partir de ces documents.
2. Annoter les entités nommées sous la forme de paires clé-valeur dans un format JSON.

Livrables attendus :

Un script ou une API capable de prendre en entrée un fichier et de renvoyer les données extraites et annotées sous forme de JSON.



NB :

Pour chaque document contenant des images (par exemple, une photo dans un passeport), celles-ci doivent être extraites et encodées en base64 avant d'être intégrées à la variable de sortie (par exemple, "photo": "data:image/jpeg;base64,/9j/4AAQSkZJRgABAQEASABIAAD/...").

Pouvoir également extraire et encoder les signatures et cachets en base64 serait un atout.

Exemple : Certificat de mariage

- Image



- Données extraite et annotation :

POST localhost:5000/api/eform/certificat-mariage

Send

200 OK4.68 s5.2 KB10 Days Ago

ParamsBodyAuthHeaders4ScriptsDocs

PreviewHeaders7CookiesTests0/0MockConsole

JSON

Preview

```
1 {
2   "text": "DUB REPUBLIQUE DEMOCRATIQUE DU CONGO PROVINCE DU LUALABA K
VILLE DEMONEMANIZA ETAT-CIVIL NUMERO:005 VOLUME: 1/2020 ACTE DE MARIAGE
L'an DEUX MILLE VINGT, le CINQUIEME jour du mois NOVEMBRE, devant nous,
FRANK MAULA KAFYOTO, Officier de l'Etat-Civil et Bourgmestre de la Commune
de Manika à Kolwezi, Province du Lualaba, REPUBLIQUE DEMOCRATIQUE DU
CONGO. Ont comparu en séance publique, le nommé FARAI CHIRODZERO né à
ZIMBABWE le DIX-HUITIEME jour du mois de MARS, l'an MIL NEUF CENT QUATRE-
VINGT ET UN, fils majeur d'âge légitime de Monsieur PARADZAI T.CHIRODZERO
et de Madame ENIA M CHIRODZERO, Etat-Civil: CELIBATAIRE, Profession:
INGENIEUR, de nationalité ZIMBABWEENNE, résidant en REPUBLIQUE
DEMOCRATIQUE DU CONGO, au numéro DEUX CENT QUATRE-VINGT-DIX-SEPT, de
l'avenue UNGU-NGANDU, quartier INDUSTRIEL, dans la commune de MANIKA à
KOLWEZI. Et Mademoiselle KAMUZE KIBAWA PATRICIA née à KOLWEZI, le ONZIEME
jour du mois de SEPTEMBRE, l'an MIL NEUF CENT QUATRE-VINGT-DEUX, fille
majeur d'âge légitime de Monsieur MPANGA KAMUZE EMMANUEL de Madame NGOIE
KIBAWA FELIE, Etat-Civil: CELIBATAIRE, Profession AGENT GENESIS, de
Nationalité: CONGOLAISE, résidant en REPUBLIQUE DEMOCRATIQUE DU CONGO, au
numéro ---, de l'avenue KIPUKA, INDUSTRIEL, dans la commune MANIKA à
KOLWEZI. Quartier W lesquels nous ont requis de procéder à la célébration
du mariage projeté entre eux et dont nous avons publié le projet formé
aux prescriptions de l'article 378 du code de la famille par une
proclamation faite sur la place PROVINCE Dubique, le VINGT-DEUXIEME jour
du mois d' OCTOBRE l'an DEUX MILLE VINGT et par une affiche 054903 apposée
à la porte de notre demeure le même jour, mois et an que dessus et nous
ont produit à cet effet: Pour le futur époux : une attestation de
```

```
1 {
2   "eform": {
3     "adresseMariage": {
4       "continent": "AFRIQUE",
5       "pays": "REPUBLIQUE DEMOCRATIQUE DU CONGO",
6       "quartier": "INDUSTRIEL",
7       "ville": "KOLWEZI"
8     },
9     "dateMariage": "05/11/2020",
10    "documentType": "certificatMariage",
11    "informationEpouse": {
12      "dateNaissance": "11/09/1982",
13      "lieuNaissance": "KOLWEZI",
14      "nom": "KAMUZE KIBAWA",
15      "parents": {
16        "mere": "NGOIE KIBAWA FELIE",
17        "pere": "MPANGA KAMUZE EMMANUEL"
18      },
19      "prenom": "PATRICIA",
20      "profession": "AGENT GENESIS",
21      "temoins": [ + 1 + ]
22    },
23    "informationEpoux": {
24      "dateNaissance": "18/03/1981",
25      "lieuNaissance": "ZIMBABWE",
26      "nom": "CHIRODZERO",
27    }
28  }
29 }
```

Beautify JSON

\$store.books[*].author