



INTÉGRATION DES BASES DE DONNÉES ET EXTRACTION DE RESSOURCES MINIERES PAR NLP ET MODELES DE LANGAGE

Encadré par : Rémy DECOUPES &
Roberto INTERDONATO
Présenté par : Aicha ZOUHAIR



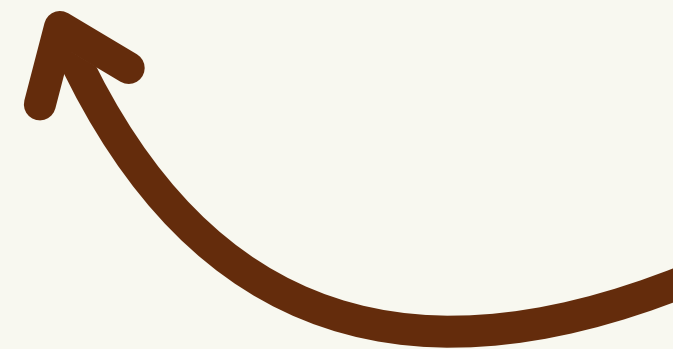
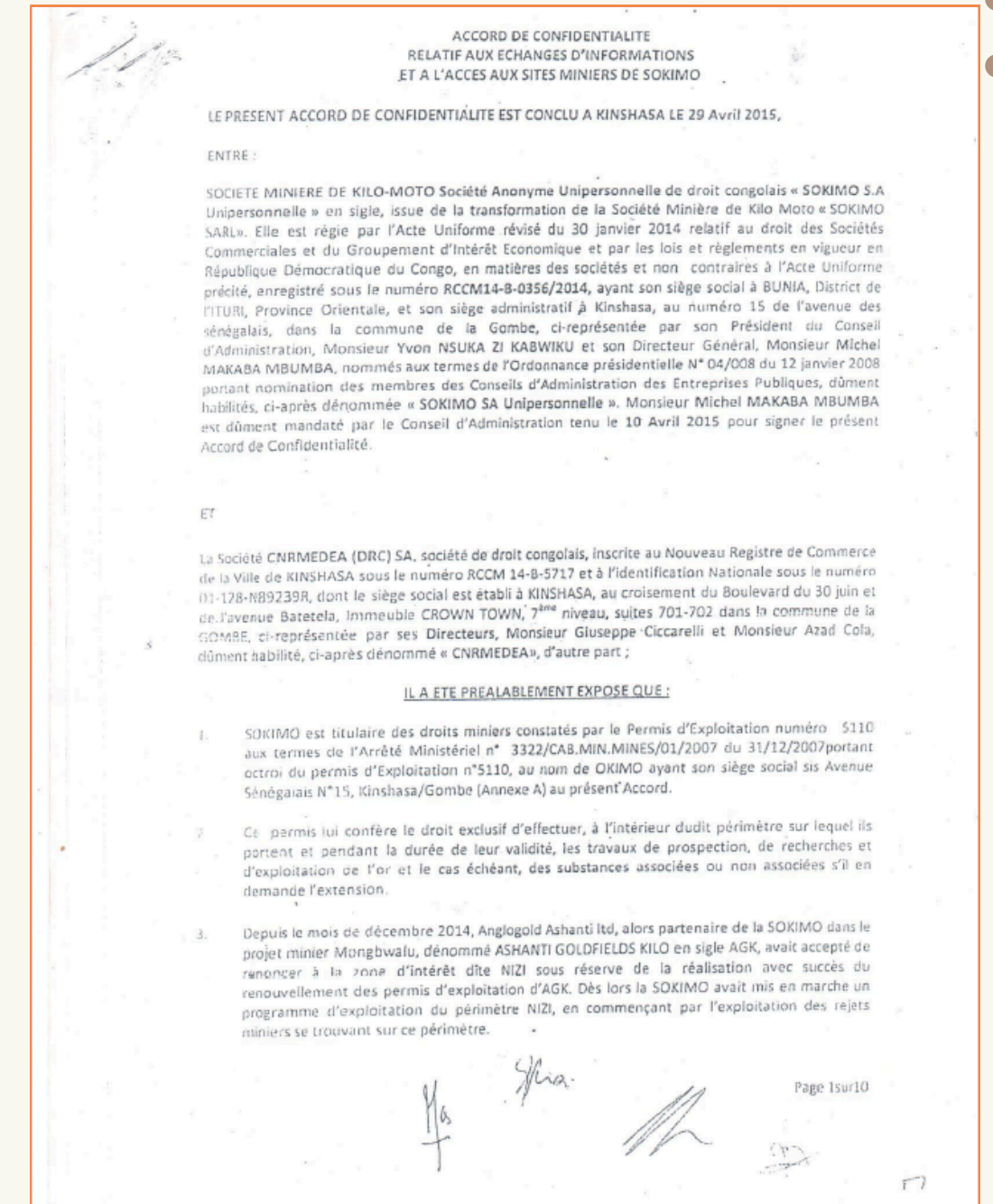
OBJECTIVE



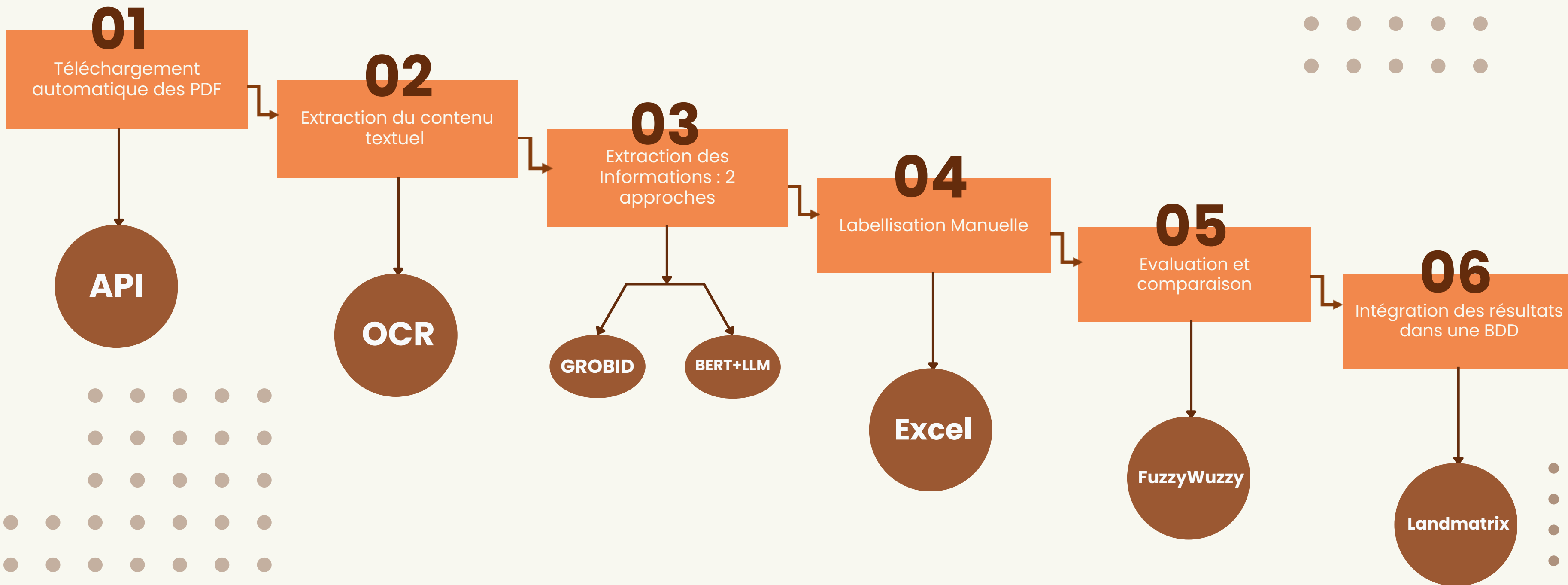
LAND MATRIX Data Observatories Resources About FAQ Contribute Deals and Investors English Login

818 Deals **Map Tables Charts**

Current intention of investment	Current negotiation status	Current implementation status	Deal size	Operating company
Mining	Concluded (Contract signed)	In operation (production)	2574.28 ha	CMOC Brasil #47294
Mining	Concluded (Contract signed)	In operation (production)	16906.22 ha	POSCO ARGENTINA S.A.U #44260
Mining	Concluded (Contract signed)	Project not started	4213.76 ha	CAMYEN S.E. - Catamarca Minera y Energetica Sociedad del Estado #44501
Mining	Concluded (Contract signed)	Project not started	4024.07 ha	CAMYEN S.E. - Catamarca Minera y Energetica Sociedad del Estado #44501
Mining	Concluded (Contract signed)	Project not started	3028.79 ha	CAMYEN S.E. - Catamarca Minera y Energetica Sociedad del Estado



PIPELINE



RESULTATS

```
{
  "file_name": "1274_Gécamines, OMG B.V., Group George Forrest S.A., JVA, 1997.pdf",
  "titre_contrat": "ACCORD DE JOINT VENTURE ENTRE",
  "pays_cible": [
    "finlande",
    "luxembourg",
    "pays-bas",
    "république démocratique du congo"
  ],
  "societe_exploitation": "Gécamines, OMG B.V., Group"
},
{
  "file_name": "1348_Gécamines, Congo Minerals SPRL, JVA, 2000.pdf",
  "titre_contrat": "CONTRAT DE CREATION DE LA SOCIETE",
  "pays_cible": [
    "république démocratique du congo"
  ],
  "societe_exploitation": "Gécamines, Congo Minerals SPRL"
},
{
  "file_name": "1349_Gécamines, Cobalt Metals Company Limited, JVA, 2000.pdf",
  "titre_contrat": "CONTRAT DE CREATION DE",
  "pays_cible": [
    "république démocratique du congo"
  ],
  "societe_exploitation": "Gécamines, Cobalt Metals Company"
},
}
```

```
... Output exceeds the size limit. Open the full output data in a text editor
Found 58 matched rows
Accuracy (titre_contrat) with fuzzy matching: 0.81
Accuracy (societe_exploitation) with fuzzy matching: 0.67

Classification Report for 'pays_cible':
```

	precision	recall	f1-score	support
afrique du sud	0.40	1.00	0.57	2
australie	1.00	1.00	1.00	1
belgique	0.60	1.00	0.75	3
british virgin island	0.00	0.00	0.00	6
canada	0.25	1.00	0.40	1
china	0.00	0.00	0.00	4
etats-unis	0.00	0.00	0.00	1
france	0.29	0.50	0.36	4
gibraltar	0.00	0.00	0.00	1
iles vierges britanniques	0.00	0.00	0.00	1
luxembourg	1.00	1.00	1.00	2
pays-bas	1.00	1.00	1.00	1
republique de coree	0.00	0.00	0.00	2
republique de finlande	0.00	0.00	0.00	1
republique democratique du congo	0.95	0.91	0.93	46
republique du congo	0.00	0.00	0.00	1
republique populaire de chine	0.00	0.00	0.00	1
republique democratique du congo	0.00	0.00	0.00	2
...				
macro avg	0.34	0.42	0.35	88
weighted avg	0.63	0.66	0.63	88
samples avg	0.80	0.73	0.74	88



THANK YOU