

BILAN D'EXTRACTION — MODÈLE QWEN

POLYCLINIQUE COTE BASQUE
SUD · FINESS 640780748 13
dossiers OGC · Contrôle 04/03/2019 ·
April 2026

1. Indicateurs globaux

12	2067	1789	278	86.6%
Dossiers analysés	Champs comparés	Champs corrects	Champs en erreur	Précision globale

2. Résultats par dossier OGC

N° OGC	Champs totaux	Champs corrects	Erreurs	Précision	Err. structurelle
OGC 358	179	158	21	88%	—
OGC 368	179	158	21	88%	—
OGC 371	178	168	10	94%	—
OGC 373	86	83	3	97%	■ Oui
OGC 385	198	149	49	75%	—
OGC 397	183	178	5	97%	—
OGC 398	181	132	49	73%	—
OGC 404	178	152	26	85%	—
OGC 411	184	164	20	89%	—
OGC 413	182	146	36	80%	—
OGC 425	161	145	16	90%	—
OGC 429	178	156	22	88%	—
TOTAL	2067	1789	278	86.6%	1 dossier(s)

3. Précision par catégorie de champ

Catégorie	Champs totaux	Champs corrects	Erreurs	Précision
DAS	189	62	127	33%
Données séjour	308	294	14	95%

DP / DR	66	43	23	65%
Données RUM	198	176	22	89%
Signataires / Dates	548	528	20	96%
Métadonnées	77	71	6	92%
Concertation (2/2)	113	87	26	77%
Éléments de preuve	408	379	29	93%
Accord / SE	55	51	4	93%
GHM / GHS	55	54	1	98%
Actes	50	44	6	88%

4. Précision par type de page

Type de page	Champs totaux	Champs corrects	Erreurs	Précision
Fiche RECUEIL	1042	839	203	81%
Éléments de preuve	852	814	38	96%
Fiche ADMIN2/2	137	108	29	79%
Fiche ADMIN1/2	36	28	8	78%

5. Patterns d'erreurs récurrents

Ce que le modèle a raté — catégorie d'erreur	Occur- rences	Sur combien de dossiers	Gravité
DR confondu avec DAS	7	7 / 13	Critique
Année mal lue (ex : 2017 au lieu de 2018)	8	6 / 13	Haute
se_coche inventé ('1' ou '4' au lieu de vide)	1	1 / 13	Haute
provenance inventé ('8' au lieu de vide)	2	2 / 13	Haute
Code acte mis dans DAS	3	3 / 13	Haute
DAS entier manquant	4	4 / 13	Critique
DAS code mauvais	7	2 / 13	Critique
JSON non parsable	1	1 / 13	Critique

6. Analyse temporelle

Résumé global

Durée totale d'extraction	1h11m18s
Durée moyenne / dossier	5m29s
Durée moyenne / page	54s
Pages traitées	78
Erreurs totales	1
Blocages rate limit (429)	0
Temps perdu en attentes 429	0s
Temps utile (hors 429)	1h11m18s

Détail par dossier

Dossier	Début	Fin	Durée	Pages	Erreurs	Blocages 429	Attente 429
23_20190411093049_00001	2026-04-20 19:36	2026-04-20 19:43	7m24s	6	0	0	—
OGC 358	2026-04-20 19:43	2026-04-20 19:47	3m49s	6	0	0	—
OGC 368	2026-04-20 19:47	2026-04-20 19:51	3m59s	6	0	0	—
OGC 371	2026-04-20 19:51	2026-04-20 19:58	6m59s	6	0	0	—
OGC 373	2026-04-20 19:58	2026-04-20 20:05	6m36s	6	1	0	—
OGC 385	2026-04-20 20:05	2026-04-20 20:10	5m00s	6	0	0	—
OGC 397	2026-04-20 20:10	2026-04-20 20:15	4m59s	6	0	0	—
OGC 398	2026-04-20 20:15	2026-04-20 20:20	5m42s	6	0	0	—
OGC 404	2026-04-20 20:20	2026-04-20 20:27	7m14s	6	0	0	—
OGC 411	2026-04-20 20:27	2026-04-20 20:33	5m01s	6	0	0	—
OGC 413	2026-04-20 20:33	2026-04-20 20:38	5m45s	6	0	0	—
OGC 425	2026-04-20 20:38	2026-04-20 20:42	4m13s	6	0	0	—
OGC 429	2026-04-20 20:42	2026-04-20 20:47	4m32s	6	0	0	—

Erreurs et blocages détaillés

OGC 373

■ Page 1 — parsing_json : JSON non parsable :

Rapport généré par comparaison automatique de extraction_ogc_raw.json vs extraction_ogc_raw_Correction.json · Périmètre : 13 dossiers OGC · Les pourcentages de précision sont calculés champ par champ.