iApp Technology · Technical Whitepaper WP-2026-23

Thai Legal Data API: Statute Search and Section Lookup Measured on NitiBench

Service version 1.0.0 · Measured 2026-10-28 · iApp Technology Co., Ltd., Bangkok, Thailand

Executive summary. The iApp Thai Legal Data API finds the Thai law section that governs a question written in plain Thai and returns any section verbatim by law and number. On the 3,729 questions of the public NitiBench-CCL benchmark it returned the governing section among its first ten results for 88.0% (95% CI 87.0–89.1%) and first for 58.4%, while searching all 2,346 laws it holds. On the 50 real tax rulings of NitiBench-Tax, long case narratives with several answer sections each, it found 22.1% of the answer sections in the first ten (CI 15.6–29.8%). Lookup returned all 3,585 benchmark sections. For sections amended since 2014, the served text carries the amending act's wording in 139 of 154 cases (the benchmark's own copy: 86); 13, mostly from the 2024 family-law amendment, are still served in the earlier wording, each with a warning. Lookup accepts only the catalogue form of a law name and Arabic section numbers.

1 Introduction

The Thai Legal Data API serves Thai statutes, Supreme Court rulings, Council of State opinions and Administrative Court decisions as structured data, over REST and as tools for AI agents. `/search` returns the sections that govern a situation described in Thai, `/section` returns one section verbatim, and `/deka/search` returns rulings, optionally only those citing a given section. We ask four questions of the production service: does search find the governing section; does lookup return the right and current text; what does the corpus hold and how recent is it; and does the public API return what the service returns.

2 Test Data

NitiBench [1] (MIT licence, revision pinned in `get_data.py`) is a Thai legal question-answering benchmark. Its *CCL* split holds 3,729 questions, each answered by exactly one section of one of 21 Thai financial and commercial laws (the Civil and Commercial Code, the Revenue Code, securities, banking, energy and company law). Its *Tax* split holds 50 real Revenue Department rulings; each is long, describes a taxpayer's facts, and is answered by 2.6 sections on average (131 in all).

Held-out status. The search components are general-purpose and were not trained on NitiBench. The service's curated 39-law set holds the same consolidated edition of the benchmark's laws, so every answer section is present: the benchmark measures finding and ranking, not coverage. The confidence floor described below was set on a separate set of hand-written probe questions (144 in the current set), of which 1 also appears verbatim in NitiBench. The grounded-answer endpoint `/ask` is not scored here, because the language model behind it was developed with NitiBench as a reference benchmark.

3 Method

Search. Each question was sent to `/search` with `top_k=20`, the maximum. A result matches when its law title and section number equal the benchmark's. The service searches all 2,346 laws and withholds results whose relevance score falls below a fixed confidence floor, so it may return fewer than 20; an answer section

that is withheld counts as a miss. For CCL we report Hit@k (the answer section is among the first k) and MRR@10; for Tax, Hit@k (at least one answer section) and Recall@k (answer sections found over all answer sections). Confidence intervals are percentile bootstraps over questions, 2,000 resamples, seed 1337.

Lookup. Every distinct answer section (3,585) was requested from `/section` and its text compared, ignoring whitespace, with the text NitiBench ships. For each section the service records as amended by an act of 2014 or later, the amending act itself was read through the same API and the wording it prescribes (“. . . shall be replaced by the following”) was checked against the served text and against the benchmark’s text. A seeded sample of 300 sections was also requested in four other common spellings.

Precedent filter. For 100 sections drawn (seed 1337) from the 400 most cited by Supreme Court rulings, `/deka/search` was asked for five rulings citing that section, and each ruling was then fetched by its case number to check that its citation list contains the section.

Two paths. The full sets were sent directly to the production service, bypassing the billing gateway, at most three requests at a time; a seeded sample went through the public API (Section 6).

4 Results

4.1 Finding the governing section

Table 1: Statute search on NitiBench (%). Failed requests: 0 (CCL), 0 (Tax).

Set	Questions	@1	@5	@10	@20	MRR@10
CCL, Hit@k	3,729	58.4	84.3	**88.0**	89.9	0.695
Tax, Hit@k	50	22.0	40.0	44.0	46.0	0.286
Tax, Recall@k	131 sections	8.4	19.8	**22.1**	22.9	

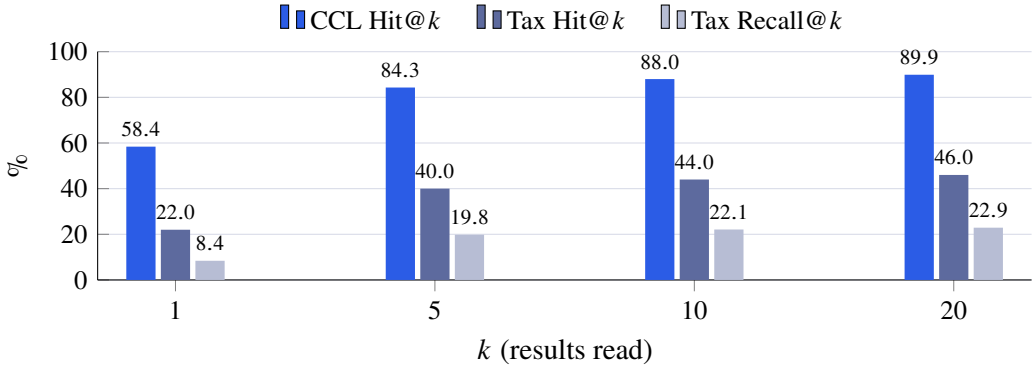


Figure 1: Share of questions (Hit) and of answer sections (Recall) found in the first k results.

The CCL answer section was found in the first ten results for 88.0% of questions (95% CI 87.0–89.1%; MRR@10 0.695, CI 0.683–0.708). The first result came from the right law for 82.6% of questions, and from a law outside the benchmark’s set for 13.2%. The service returned fewer than 20 results for 67.1% of questions (13.1 on average; 33 returned none); of the 378 questions whose answer was not returned at all, 283 had results withheld by the confidence floor. Tax rulings are long and cite several provisions; the service returned 4.1 results on average and none for 6 of 50, and found at least one answer section in the first ten for 44.0% (CI 32.0–58.0%).

Context. The benchmark’s authors report, for 8 retrievers each searching only the 35 benchmark laws, CCL Hit@10 of 71.5–93.8% (Hit@1 48.1–73.5%) and Tax Hit@10 of 22.0–84.0% with Recall@10 of

8.5–47.2% [1]. Those figures are not directly comparable: the API searches 2,346 laws rather than 35, and withholds low-confidence results instead of always returning ten.

Table 2: CCL by law (%), ordered by number of questions.

Law	Questions	Hit@1	Hit@10	MRR@10
Civil and Commercial Code	1,617	60.4	87.0	0.706
Revenue Code	484	56.6	87.0	0.678
Securities and Exchange Act 1992	294	62.2	92.2	0.728
Public Limited Companies Act 1992	186	60.2	94.6	0.731
Financial Institutions Business Act 2008	165	58.2	90.3	0.697
Energy Industry Act 2007	153	47.1	78.4	0.579
Petroleum Income Tax Act 1971	110	43.6	80.0	0.550
Derivatives Act 2003	109	54.1	89.9	0.682
Digital Asset Businesses Decree 2018	78	57.7	91.0	0.690
Business Security Act 2015	74	63.5	95.9	0.769
Capital Market Trust Act 2007	73	58.9	91.8	0.718
Accounting Profession Act 2004	64	75.0	96.9	0.833
Energy Conservation Promotion Act 1992	58	46.6	74.1	0.562
Partnership and Company Offences Act 1956	48	33.3	85.4	0.545
Trade Associations Act 1966	40	62.5	95.0	0.772
Accounting Act 2000	38	57.9	84.2	0.668
Foreign Business Act 1999	36	75.0	97.2	0.853
Provident Fund Act 1987	36	61.1	91.7	0.739
Securitisation Decree 1997	33	60.6	87.9	0.705
Chambers of Commerce Act 1966	22	40.9	90.9	0.594
Commercial Registration Act 1956	11	54.5	100.0	0.758

4.2 Section lookup and currency of the text

All 3,585 answer sections were found (100.0%). Their text is identical to the benchmark's for 3,507 (97.8%); each of the other 78 carries an amendment note naming a later amending act, which the benchmark's copy lacks. The amendment check shows which copy is current. Of the 154 sections amended from 2014 onward whose new wording could be read from the amending act, the served text carries that wording in 139 (90.3%) and the benchmark's in 86 (55.8%); in 2 more the served text is newer still, reflecting a later conversion of criminal fines into administrative fines. The remaining 13 are served in the earlier wording: 12 sections of the Civil and Commercial Code amended in 2024 to make engagement and marriage gender-neutral (s. 1443, 1446, 1450, 1453, 1470, 1487, 1499, 1533, 1538, 1543, 1544, 1545; 41 of the 53 tested sections of that amendment are current), and Digital Asset Businesses Decree 2018 s. 26 (amended 2025). All 13 of these responses carry a warning that names the amending act.

Table 3: Lookup of 300 seeded sections written in other common ways. "Suggests" counts misses whose suggestions include the right law; "Reported unknown law" counts misses answered with `law_not_found`.

How the request was written	Requests	Found	Suggests	Reported unknown law
Catalogue name, Arabic numerals (documented)	300	300	0	0
Section number in Thai numerals	300	0	300	300
Section number with the word for "section"	300	0	300	300
Abbreviated law name	300	0	101	300
Act name without its year	142	0	142	142

Only the documented form works. A section number in Thai numerals, as printed in the Royal Gazette, or prefixed with the word for "section", is not found, and the response reports the law as unknown although the law is held. Abbreviated code names, the usual way to cite the Civil and Commercial Code, return no

suggestion (0 of 158); abbreviated act names suggest the right law in 101 of 142.

4.3 Precedent filter

The 100 filtered searches returned 483 rulings (3 searches returned none). Of the 471 that could be fetched by case number, 443 (94.1%) list the filtered section among their citations. 12 rulings could not be fetched: all of them carry joined case numbers such as "1579-1580/2558", which `/deka/search` returns but `/deka/{case_id}` does not accept in either documented form; the public API behaves the same (12 of 12). In 37 searches the same ruling appeared more than once in the five results (43 repeated entries), because a ruling with several headnotes is indexed once per headnote.

4.4 Corpus coverage and freshness

Table 4: What the API reports holding, read from its own `/meta`, `/laws` and `/facets` on 2026-10-28.

Source	Records	Span and freshness
Statutes	2,346 laws	68,237 operative sections; latest law dated 2026; 200 laws dated 2024 or later. In-force status: 1,087 laws confirmed in force by the national legal register, 275 repealed, 982 unverified.
Supreme Court	133,159	120,216 judgments plus orders and rulings, 1920–2026; 1,079 from 2024 on, 51 from 2026; 99.7% linked to the sections they cite.
Council of State	6,534	1923–2026; 6,525 with full text; 25 from 2026.
Administrative Court	128,361	16 courts, 1981–2026; 49,862 with full text; latest decision dated 2026-06-10.

5 Discussion

Search is strong on questions about one provision and weaker on case facts. A short question usually finds its section on the first page of results although the search spans every law held. A long tax narrative that turns on several sections gets at least one of them, rarely all; a caller should search per issue rather than with the whole narrative.

The confidence floor trades recall for fewer distractors. A caller, often an AI agent, receives fewer plausible but wrong sections. Its cost cannot be isolated from outside the service, but 283 of the 378 CCL questions whose answer was not returned received fewer than 20 results, so lower-scored candidates had been withheld.

The statute text is more current than the benchmark's. Where the two disagree, the service holds the amended wording in most cases; the exceptions are concentrated in one amendment and each is flagged, so a caller must read the warning until the text is updated.

Lookup is strict. Thai numerals and abbreviated code names are how Thai lawyers write citations; both fail, and the "unknown law" status misleads.

6 Production Performance

A seeded sample of 250 search questions and 50 section lookups went through `https://api.iapp.co.th/v3/store/data/thai-legal` with an ordinary API key, from two concurrent clients paced below the route's per-minute rate limit. None of the requests failed (0 search, 0 lookup). Lookups returned text identical to the service's for 50 of 50. For search, the public API gave the same first result as the earlier full

run for 241 of 250 questions and the same answer-in-top-ten outcome for 250, but an identical ordered list for only 86. The same questions sent to the service directly a second time agree with the first run to the same degree (243, 249 and 88 of 250): relevance scores vary slightly between identical requests (median difference 0.0003, largest 0.052), which reorders near-tied results. The gateway adds no difference of its own.

Table 5: Latency through the public API, measured at the client, request time only.

Endpoint	Requests	p50	p90	p95
`/search` (s)	250	0.88	1.12	1.20
`/section` (ms)	50	69	76	78

The benchmark made 624 public calls, counted from the account's usage records; the endpoints were free during the launch promotion, so 0 IC was charged.

7 Limitations

- **One domain.** NitiBench covers financial and commercial law; criminal, labour, land and administrative law, which the API also serves, are not measured. Matching is exact: a closely related section counts as a miss.
- **Amendment check.** It tests only amendments the service's own register records and whose wording the amending act states in full (5 recorded amendments could not be tested); amendments the register lacks are not detected.
- **Not measured.** The grounded-answer endpoint, ruling search without a citation filter, and the Council of State and Administrative Court searches.

8 Reproducibility

The test data are public (NitiBench, MIT licence). `get_data.py` fetches the pinned revision, `bench.py` sends every item and stores each response, and `gen_results.py` computes every figure in this paper from those files. The per-question results and the scoring code are available from iApp Technology on request: sale@iapp.co.th.

References

[1] P. Akarajaradwong, P. Pothavorn, C. Chaksangchaichot, P. Tasawong, T. Nopparatbundit, S. Nutanong. *NitiBench: A Comprehensive Study of LLM Frameworks' Capabilities for Thai Legal Question Answering*. arXiv:2502.10868, 2025. Dataset: https://huggingface.co/datasets/VISAI-AI/nitibench

About iApp Technology. iApp Technology Co., Ltd. is a Thai artificial-intelligence company providing OCR, document understanding, speech, identity verification and large-language-model services for the Thai language.
Web: iapp.co.th · E-mail: sale@iapp.co.th · Tel: +66 2 124 4041