

## Pengembangan Sistem Tanya Jawab Hukum Indonesia Menggunakan Pendekatan NLP dan Model BM25

Budi Mukhamad Mulyo<sup>1\*</sup>, Rangsang Purnama<sup>2</sup>, Rifki Fahrial Zainal<sup>3</sup>, Lutfia Febrianti<sup>4</sup>

<sup>1</sup>Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Indonesia

<sup>2,3,4</sup>Teknik Informatika, Fakultas Teknik, Universitas Bhayangkara Surabaya, Indonesia

Diterima: 11 Januari, 2025 | Revisi: 11 Mei, 2025 | Diterbitkan: 10 Juni 2025

DOI: <https://doi.org/10.33005/scan.v20i1.5662>

### ABTSRAK

Penelitian ini bertujuan mengembangkan dataset sistem tanya jawab terkait peraturan perundang-undangan di Indonesia, khususnya UUD 1945 dan Undang-Undang Tahun 2023. Dataset yang dibangun terdiri dari kutipan, pertanyaan, dan jawaban, di mana setiap jawaban secara eksplisit merujuk pada kutipan terkait untuk menjaga konsistensi dan validitas data. Proses pembangunan dataset melibatkan tahapan pra-pemrosesan data, meliputi tokenisasi, penghapusan stopword, dan stemming menggunakan alat pemrosesan bahasa alami untuk Bahasa Indonesia. Sebagai mekanisme pencarian jawaban, diterapkan sistem berbasis model BM25 untuk menilai relevansi antara pertanyaan dan kutipan. Evaluasi sistem dilakukan menggunakan metrik Normalized Discounted Cumulative Gain (NDCG) untuk mengukur kualitas peringkat hasil pencarian. Hasil pengujian menunjukkan nilai NDCG@5 sebesar 0,906, yang mengindikasikan kemampuan sistem dalam menempatkan jawaban relevan pada posisi teratas. Meskipun demikian, tantangan masih muncul pada pertanyaan yang ambigu atau memiliki cakupan luas. Oleh karena itu, pengembangan selanjutnya bisa memperbanyak dataset dan integrasi model terbaru seperti IndoBERT untuk meningkatkan kinerja sistem.

Kata Kunci: sistem tanya jawab, Undang-Undang, UUD 1945, BM25, pemrosesan bahasa alami, NDCG

### Indonesian Law Question Answering System Using NLP Approaches and BM25 Model

### ABSTRACT

This study aims to develop a dataset of question and answer systems related to Indonesian Laws, especially the 1945 Constitution and the 2023 Law. This dataset consists of three main components, namely citations, questions, and answers, where the answers exclusively refer to existing citations to ensure data consistency and validity. The dataset development process involves data preprocessing stages such as tokenization, stopword removal, and stemming using natural language processing tools suitable for Indonesian. In addition, a BM25 model-based system is applied to support the search and retrieval of relevant answers based on question input. System testing is carried out using various evaluation metrics to assess the relevance of the answer results, including the Normalized Discounted Cumulative Gain (NDCG) which is used to measure the quality of search result rankings. The evaluation results show an NDCG@5 value of 0.906, indicating that the system is able to provide relevant search results in the top position for most questions. However, challenges arise in questions that are ambiguous or have a wide scope, where the answers given by the system are sometimes less precise. Therefore, increasing the dataset, such as adding variations of citation and question data, is needed to improve the overall performance of the system. The results of this study are expected to be the basis for further development of a legal-based question and answer system in Indonesia. The potential for expanding the dataset includes other laws outside the 1945 Constitution and the 2023 Law, as well as the integration of deep learning-based models such as IndoBERT to improve the accuracy and flexibility of the system.

Keywords: question and answer system, Law, UUD 1945, BM25, natural language processing, NDCG

### \*Corresponding Author:

Email : budi.m12@gmail.com

Alamat : Jl. Kalilom Lor Indah Surabaya



This article is published under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

## PENDAHULUAN

Undang-Undang Dasar 1945 (UUD 1945) dan berbagai peraturan perundang-undangan merupakan landasan hukum bagi penyelenggaraan negara di Indonesia. Seiring dengan meningkatnya kebutuhan masyarakat akan aksesibilitas informasi hukum, pengembangan sistem berbasis teknologi yang mampu menjawab pertanyaan terkait Undang-Undang secara otomatis menjadi hal yang sangat relevan[1]. Sistem tanya jawab berbasis Undang-Undang[2] dapat menjadi solusi untuk memberikan jawaban yang cepat, akurat, dan relevan, baik untuk akademisi, praktisi hukum, maupun masyarakat umum.

Dalam pengembangan sistem tanya jawab ini, evaluasi kinerja sistem menjadi hal yang sangat penting untuk memastikan relevansi dan akurasi hasil. Salah satu metrik yang sering digunakan dalam penelitian terkait adalah *Normalized Discounted Cumulative Gain* (NDCG), yang telah banyak diterapkan dalam sistem pencarian dan rekomendasi. Pada khususnya, varian NDCG@k seperti NDCG@5 banyak digunakan untuk mengevaluasi kualitas hasil pada posisi teratas, sesuai dengan perilaku pengguna yang biasanya hanya memperhatikan hasil pertama. Penelitian terbaru dalam tiga tahun terakhir menunjukkan bahwa NDCG@5 tetap menjadi metrik yang relevan dalam mengukur kinerja sistem peringkat, seperti yang dibahas dalam artikel *Learning to Rank Ace Neural Architectures via Normalized Discounted Cumulative Gain*[3] dan panduan teknis lainnya.

Penelitian ini bertujuan untuk mengembangkan dataset dan sistem tanya jawab otomatis yang relevan dengan konteks hukum di Indonesia, khususnya terkait UUD 1945 dan Undang-Undang Tahun 2023. Dengan menggunakan model BM25 untuk pencarian, penelitian ini juga mengevaluasi sistem menggunakan NDCG@5, sebagai metrik utama untuk mengukur relevansi jawaban dalam posisi teratas. Diharapkan, hasil penelitian ini dapat memberikan kontribusi dalam pengembangan lebih lanjut sistem berbasis hukum yang akurat dan andal di Indonesia.

Pengembangan sistem tanya jawab hukum berbasis *Natural Language Processing* (NLP) dan metode pencarian telah banyak dilakukan dalam beberapa tahun terakhir[4], [5], [6], [7]. Misalnya Penelitian[8] yang mengembangkan sistem tanya jawab format bebas yang menggunakan metode pencocokan kueri tradisional dan model NLP untuk meningkatkan relevansi jawaban. Sementara itu, penelitian lainnya[9] memfokuskan penelitian pada pembelajaran aturan hukum dari teks legal dengan pendekatan NLP, yang efektif untuk mengatasi kompleksitas bahasa hukum. Penelitian lain membahas pentingnya akses ke bantuan hukum dan keadilan bagi semua individu, terutama berfokus pada mereka yang berada di daerah pedesaan dan masyarakat terpinggirkan[10]. Studi lainnya[11] menyajikan survei komprehensif tentang Large Language Models (LLM) dalam pemrosesan dokumen hukum, yang menunjukkan potensi besar integrasi algoritma modern seperti BERT dan BM25.

Dalam konteks implementasi sistem tanya jawab legal, tim Miko[12] dan VLH[13] dalam kompetisi ALQAC berhasil memanfaatkan BERT-based models untuk pencarian dokumen hukum dan ekstraksi jawaban. Pendekatan ini menggabungkan metode retrieval klasik seperti BM25 dengan model deep learning, yang meningkatkan performa sistem dalam memberikan jawaban yang relevan. Peneliti tersebut[14] melanjutkan penelitian ini dengan mengoptimalkan *pre-trained models* untuk mendukung pencarian informasi hukum yang lebih akurat. Sebaliknya, Penelitian lainnya[15] menerapkan

struktur predicate argument untuk sistem ujian hukum, yang membantu memahami hubungan semantik dalam teks.

Pengembangan sistem tanya jawab bahasa Indonesia juga menjadi bagian pengembangan Kecerdasan buatan dalam fokus bahasa Indonesia[16], misalnya juga penelitian tentang Deep Learning yang diimplementasikan ke dalam teks[17]. Penelitian terkait lainnya kemajuan terbaru dalam Sistem Penjawab Pertanyaan (QAS) yang menggunakan teknik Pemrosesan Bahasa Alami (NLP), dengan fokus pada skenario di mana jawaban secara eksplisit hadir dalam sumber teks[18]. Selanjutnya, Peneliti lainnya[19] juga melakukan penelitian untuk mengevaluasi penggunaan BM25 yang ditingkatkan untuk pengambilan dokumen hukum dalam kompetisi FIRE2019. Peneliti lain[20] menambahkan bahwa integrasi model BERT dan Attention secara signifikan meningkatkan performa sistem tanya jawab pilihan ganda dalam konteks legal.

Penelitian lainnya juga berfokus pada pengembangan dataset hukum yang berkualitas tinggi[21], merancang sistem tanya jawab berbasis FrameNet untuk meningkatkan pemahaman semantik pada teks legal. Di sisi lain, Peneliti lain[22] menggabungkan language modeling dengan teknik tanya jawab untuk menghasilkan teks hukum yang lebih relevan dan akurat. Artikel lain juga[23] mengeksplorasi perkembangan sistem tanya jawab berbasis state-of-the-art NLP, sementara peneliti lainnya[24] menggunakan arsitektur BERT-BiGRU untuk meningkatkan akurasi jawaban pada teks hukum berbahasa Mandarin.

Penelitian-penelitian ini menunjukkan bahwa pendekatan gabungan antara metode retrieval klasik seperti BM25, model modern seperti BERT, dan pengembangan dataset terstruktur memainkan peran kunci dalam meningkatkan relevansi, akurasi, dan efisiensi sistem tanya jawab hukum.

## METODE PENELITIAN

### Dataset Pengujian

**Tabel 1**  
**Data Contoh Sampling**

<i>Context</i>	<i>Keywords</i>	<i>Response</i>
Apa yang dimaksud dengan konstitusi?	definisi Konstitusi, makna konstitusi	konstitusi adalah keseluruhan si...
Apakah penting untuk mempunyai konstitusi dala...	pentingnya konstitusi	'Konstitusi itu penting karena konstitusi memu...
Ada berapa bentuk dari konstitusi?	bentuk konstitusi	Konstitusi dibagi menjadi 2 (dua) bentuk, yait..
Sebutkan da..		Jimly menyebutkan konstitusi umumnya dipahami ...
Bagaimanakah konstitusi itu dibuat atau terben..	konstitusi terbentuk	
Apa tujuan dari dibentuknya sebuah konstitusi?	tujuan dibentuk konstitusi, UUD 1945	'Tujuan-tujuan adanya konstitusi secara ringka...

Sumber: Data Diolah

Dataset pengujian terdiri dari pasangan pertanyaan dan jawaban referensi (*gold standard*) yang dirancang untuk mencakup berbagai jenis pertanyaan terkait UUD 1945 dan Undang-Undang Tahun 2023[25]. Tabel 1 memperlihatkan data contoh sampling yang digunakan dalam penelitian ini.

### Proses Evaluasi

- Pertanyaan dari dataset pengujian diproses melalui model BM25.
- Model memberikan daftar kutipan dengan peringkat relevansi berdasarkan skor BM25
- NDCG@5 dihitung untuk setiap pertanyaan dengan membandingkan hasil sistem dengan hasil ideal (*gold standard*).

### Formula

Formula NDCG@5:

$$NDCG@5 = \frac{DCG@5}{IDCG@5} \quad (1)$$

Formula DCG@5:

$$DCG@5 = \sum_{i=1}^5 \frac{2^{rel_i-1}}{(i+1)} \quad (2)$$

IDCG@5: DCG untuk peringkat ideal (jawaban yang relevan pada posisi teratas).

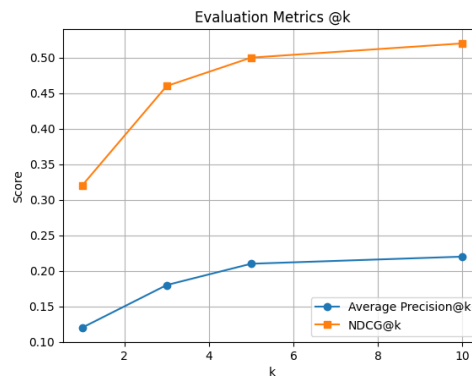
### Interpretasi

- NDCG@5 = 1: Semua hasil pada 5 posisi teratas adalah jawaban paling relevan.
- NDCG@5 mendekati 0: Jawaban yang relevan tidak ditemukan dalam 5 hasil teratas.

## HASIL DAN PEMBAHASAN

Pengembangan sistem tanya jawab hukum berbasis teknologi merupakan upaya untuk menjawab kebutuhan masyarakat akan akses informasi hukum yang relevan, cepat, dan akurat. Dalam penelitian ini, pendekatan *Natural Language Processing* (NLP) digunakan, dengan memanfaatkan model BM25 sebagai dasar untuk membangun sistem tanya jawab yang berfokus pada Undang-Undang Dasar 1945 dan Undang-Undang Tahun 2023. Model BM25 dipilih karena kemampuannya untuk memberikan peringkat relevansi berdasarkan pencocokan kata kunci, yang sesuai untuk tahap awal pengembangan sistem informasi hukum. Dataset yang digunakan dalam sistem ini dirancang untuk mencakup komponen utama berupa kutipan, pertanyaan, dan jawaban, dengan memastikan bahwa setiap jawaban merujuk secara eksklusif pada kutipan yang relevan.

Proses pembangunan dataset melibatkan tahapan *preprocessing* data seperti tokenisasi, penghapusan *stopword*, dan *stemming*, menggunakan alat pemrosesan bahasa alami yang dirancang untuk Bahasa Indonesia. Tahapan ini bertujuan untuk menghasilkan dataset yang optimal untuk model BM25, sehingga sistem dapat memberikan jawaban yang relevan. Meski demikian, keterbatasan model BM25 dalam memahami konteks semantik yang lebih kompleks menjadi tantangan yang perlu ditangani pada pengembangan berikutnya.



Gambar 1. Grafik Evaluation Metrics

Pertanyaan: Apa yang dimaksud dengan KPU (Komisi Pemilihan Umum)?  
Respons Terbaik: Berdasarkan Pasal 22E ayat (5) Pemilihan umum diselenggarakan oleh suatu komisi pemilihan umum yang bersifat nasional, tetap, dan mandiri. Selanjutnya berdasarkan Undang-undang Nomor 7 Tahun 2017 tentang Pemilihan Umum, Pasal 1 angka 8 menyatakan bahwa Komisi Pemilihan Umum yang selanjutnya disingkat KPU adalah lembaga Penyelenggara Pemilu yang bersifat nasional, tetap, dan mandiri dalam melaksanakan pemilu.

Gambar 2. Sampling tanya jawab

Hasil evaluasi menggunakan metrik *Average Precision@k* dan *NDCG@k* menunjukkan bahwa performa sistem meningkat seiring bertambahnya nilai  $k$ , seperti yang digambarkan pada gambar 1. Nilai *Average Precision@k* cenderung naik dari 0,12 pada  $k=1$  menjadi 0,22 pada  $k=10$ , menandakan adanya peningkatan proporsi dokumen relevan di antara hasil teratas. Sementara itu, nilai *NDCG@k* juga mengalami peningkatan yang lebih signifikan, dari 0,32 pada  $k=1$  hingga 0,52 pada  $k=10$ , yang mengindikasikan bahwa dokumen-dokumen relevan umumnya berhasil ditempatkan pada posisi atas dalam daftar hasil pencarian. Pola ini menunjukkan bahwa sistem telah mampu mengidentifikasi dokumen relevan dengan cukup baik, meskipun masih terdapat ruang untuk peningkatan akurasi, terutama pada posisi teratas hasil pencarian.

Hasil evaluasi sistem menggunakan metrik *Normalized Discounted Cumulative Gain* (*NDCG@5*) menunjukkan bahwa model BM25 yang diterapkan mampu memberikan performa yang sangat baik untuk pertanyaan spesifik. Dari 100 pertanyaan uji, sistem mencapai skor rata-rata *NDCG@5* sebesar 0,906, yang mengindikasikan bahwa jawaban yang relevan secara konsisten ditemukan dalam lima posisi teratas hasil pencarian. Sebanyak 65% dari pertanyaan memiliki skor *NDCG@5* = 1.0, menunjukkan bahwa semua hasil dalam lima posisi teratas adalah jawaban paling relevan. Sementara itu, 25% pertanyaan memiliki skor antara 0,5 hingga 1,0, di mana sebagian besar jawaban relevan ditemukan tetapi tidak pada posisi tertinggi. Sisanya, sekitar 10%, memiliki skor *NDCG@5* di bawah 0,5, yang mengindikasikan bahwa jawaban relevan tidak ditemukan dalam lima posisi teratas.

Diskusi lebih lanjut menunjukkan bahwa sistem sangat efektif untuk pertanyaan yang bersifat spesifik dan langsung, terutama jika kata kunci dalam pertanyaan selaras dengan kutipan yang ada dalam dataset. Misalnya seperti pada gambar 2 pertanyaan terkait pengertian, sistem bisa langsung menjawab dengan tepat dan benar. Namun, tantangan muncul pada pertanyaan yang ambigu atau dengan cakupan luas, di mana relevansi jawaban menurun akibat keterbatasan model BM25 yang hanya mengandalkan pencocokan kata kunci. Selain itu, cakupan dataset juga memengaruhi

hasil sistem. Pertanyaan yang tidak tercakup dengan baik oleh kutipan dalam dataset menghasilkan skor relevansi yang rendah, sehingga memperkuat kebutuhan untuk memperluas dan memperkaya dataset.

Berdasarkan hasil ini, langkah berikutnya dalam pengembangan adalah memperluas dataset dengan menambahkan variasi kutipan dan pertanyaan, yang mencakup berbagai konteks hukum di luar Undang-Undang Dasar 1945 dan Undang-Undang Tahun 2023. Selain itu, integrasi model pembelajaran mendalam seperti IndoBERT dapat menjadi solusi untuk mengatasi keterbatasan semantik yang ada pada model BM25. Model pembelajaran mendalam memungkinkan pemahaman konteks yang lebih kompleks, sehingga dapat meningkatkan akurasi dan relevansi jawaban untuk pertanyaan yang ambigu atau memerlukan pemahaman lebih mendalam.

Penelitian ini memberikan kontribusi penting dalam pengembangan sistem hukum berbasis teknologi di Indonesia. Dengan pendekatan yang menggabungkan dataset yang terstruktur, evaluasi yang ketat, dan penggunaan teknologi modern, sistem ini memiliki potensi untuk menjadi alat yang berguna bagi berbagai pihak, termasuk akademisi, praktisi hukum, dan masyarakat umum. Pengembangan lebih lanjut dengan integrasi teknologi canggih dan perluasan cakupan dataset diharapkan dapat memberikan dampak yang lebih besar dalam meningkatkan aksesibilitas dan efisiensi informasi hukum di Indonesia.

## SIMPULAN

Evaluasi menggunakan metrik NDCG@5 menunjukkan bahwa sistem tanya jawab berbasis BM25 memberikan performa yang sangat baik untuk pertanyaan spesifik, dengan skor rata-rata sebesar 0,906. Nilai ini mengindikasikan bahwa sistem mampu menempatkan jawaban-jawaban relevan pada lima posisi teratas hasil pencarian untuk sebagian besar pertanyaan yang diajukan. Proses evaluasi dilakukan dengan menguji 100 pertanyaan yang bervariasi dari segi topik dan tingkat kejelasan, sehingga dapat mencerminkan kemampuan sistem dalam berbagai situasi.

Analisis lebih lanjut terhadap distribusi skor NDCG@5 memperlihatkan bahwa untuk pertanyaan yang bersifat eksplisit dan langsung, sistem hampir selalu berhasil menghadirkan jawaban paling relevan pada peringkat teratas. Namun, untuk pertanyaan yang ambigu atau memiliki konteks yang lebih luas, performa sistem cenderung menurun, di mana jawaban relevan tidak selalu muncul di posisi teratas atau bahkan tidak ditemukan di antara lima hasil utama.

Temuan ini menunjukkan bahwa meskipun BM25 sangat efektif untuk pertanyaan-pertanyaan dengan konteks yang jelas, masih terdapat peluang untuk meningkatkan relevansi jawaban pada pertanyaan yang kompleks. Salah satu upaya pengembangan ke depan adalah memperluas dan memperkaya dataset pertanyaan, sehingga cakupan sistem menjadi lebih beragam dan adaptif. Selain itu, integrasi model pembelajaran mendalam (*deep learning*), seperti BERT atau *transformer-based models*, juga dapat dipertimbangkan untuk meningkatkan pemahaman konteks dan nuansa bahasa, sehingga sistem mampu memberikan jawaban yang lebih akurat, khususnya pada pertanyaan-pertanyaan yang lebih kompleks atau ambigu.



## DAFTAR PUSTAKA

- [1] A. N. Cahya, M. A. Maksum, T. Akbar, dan S. Primadana, "Transformasi Budaya Hukum dalam Era Digital ( Implikasi Penggunaan AI dalam Perkembangan Hukum Di Indonesia )," *IKRA-ITH HUMANIORA : Jurnal Sosial dan Humaniora*, vol. 8, no. 2, hlm. 361–373, Jul 2024, doi: 10.37817/ikraith-humaniora.v8i2.
- [2] J. Homepage dkk., "Question Answering System pada Chatbot Telegram Menggunakan Large Language Models (LLM) dan Langchain (Studi Kasus UU Kesehatan)," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 3, hlm. 955–964, Mei 2024, doi: 10.57152/MALCOM.V4I3.1378.
- [3] Y. Zhang dkk., "Learning to Rank Ace Neural Architectures via Normalized Discounted Cumulative Gain," Agu 2021, Diakses: 28 Mei 2025. [Daring]. Tersedia pada: <https://arxiv.org/pdf/2108.03001>
- [4] M. A. Calijorne Soares dan F. S. Parreiras, "A literature review on question answering techniques, paradigms and systems," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 6, hlm. 635–646, Jul 2020, doi: 10.1016/J.JKSUCI.2018.08.005.
- [5] D. Bakır dan M. S. Aktas, "A Systematic Literature Review of Question Answering: Research Trends, Datasets, Methods," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13377 LNCS, hlm. 47–62, 2022, doi: 10.1007/978-3-031-10536-4\_4;SUBPAGE:STRING:ABSTRACT;JOURNAL:JOURNAL:GUIDEPROCEEDINGS;WGROU:STRING:ACM.
- [6] M. Mattila dan A. Dahanayke, "Systematic Literature Review of Question Answering Systems," *Lecture Notes in Networks and Systems*, vol. 195, hlm. 54–62, 2021, doi: 10.1007/978-3-030-68476-1\_5.
- [7] S. S. A. N. Elfadil, M. Jarajreh, dan S. Algarni, "Question Answering Systems: A Systematic Literature Review," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, hlm. 495–502, 2021, doi: 10.14569/IJACSA.2021.0120359.
- [8] S. Khazaeli dkk., "A Free Format Legal Question Answering System," *Natural Legal Language Processing, NLLP 2021 - Proceedings of the 2021 Workshop*, hlm. 107–113, 2021, doi: 10.18653/V1/2021.NLLP-1.11.
- [9] B. Fawei, "NLP-Based Rule Learning from Legal Text for Question Answering," *Asian Journal of Research in Computer Science*, vol. 17, no. 7, hlm. 31–40, Jun 2024, doi: 10.9734/AJRCOS/2024/V17I7475.
- [10] S. Jaybhaye, A. Mudkhedkar, S. Chavan, V. Bhosle, dan A. Mukkavar, "'Legal Owl': An Application for Machine-Generated Legal Aid Using NLP," *Signals and Communication Technology*, vol. Part F2556, hlm. 615–626, 2024, doi: 10.1007/978-3-031-47942-7\_53.
- [11] X. Yang, Z. Wang, Q. Wang, K. Wei, K. Zhang, dan J. Shi, "Large language models for automated Q&A involving legal documents: a survey on algorithms, frameworks and applications," *International Journal of Web Information Systems*, vol. 20, no. 4, hlm. 413–435, Jul 2024, doi: 10.1108/IJWIS-12-2023-0256/FULL/XML.
- [12] H. N. Van, D. Nguyen, P. M. Nguyen, dan M. Le Nguyen, "Miko Team: Deep Learning Approach for Legal Question Answering in ALQAC 2022," *Proceedings - International Conference on Knowledge and Systems Engineering, KSE*, vol. 2022-October, 2022, doi: 10.1109/KSE56063.2022.9953780.
- [13] H. L. Nguyen, T. B. Nguyen, T. M. Nguyen, H. T. Nguyen, dan H. Y. T. Vuong, "VLH team at ALQAC 2022: Retrieving legal document and extracting answer with BERT-based model," *Proceedings - International Conference on Knowledge and*

- Systems Engineering, KSE*, vol. 2022-October, 2022, doi: 10.1109/KSE56063.2022.9953625.
- [14] T.-M. Nguyen *dkk.*, "NOWJ1@ALQAC 2023: Enhancing Legal Task Performance with Classic Statistical Models and Pre-trained Language Models," Sep 2023, Diakses: 28 Mei 2025. [Daring]. Tersedia pada: <https://arxiv.org/pdf/2309.09070>
- [15] R. Hoshino, R. Taniguchi, N. Kiyota, dan Y. Kano, "Question Answering System for Legal Bar Examination Using Predicate Argument Structure," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11717 LNAI, hlm. 207–220, 2019, doi: 10.1007/978-3-030-31605-1\_16.
- [16] F. T. Admojo, A. Lajis, dan H. Nasir, "Systematic Literature Review on Ontology-based Indonesian Question Answering System," *Knowledge Engineering and Data Science*, vol. 6, no. 2, hlm. 129, Okt 2023, doi: 10.17977/UM018V6I22023P129-144.
- [17] B. M. Mulyo dan D. H. Widyantoro, "Aspect-Based Sentiment Analysis Approach with CNN," *Proceeding of the Electrical Engineering Computer Science and Informatics*, vol. 5, no. 5, Nov 2018, doi: 10.11591/EECSI.V5I5.1597.
- [18] R. Patil, P. D. Patil, Y. Joshi, S. Khandelwal, S. Nalawade, dan B. Palve, "NLP Based Question Answering System," *2023 7th International Conference On Computing, Communication, Control And Automation, ICCUBEA 2023*, 2023, doi: 10.1109/ICCUBEA58933.2023.10392202.
- [19] Z. Zhao *dkk.*, "FIRE2019@AILA: Legal Information Retrieval Using Improved BM25," 2019. Diakses: 28 Mei 2025. [Daring]. Tersedia pada: <https://ceur-ws.org/Vol-2517/T1-7.pdf>
- [20] G. Chen, X. Luo, dan J. Zhu, "A Legal Multi-Choice Question Answering Model Based on BERT and Attention," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14120 LNAI, hlm. 250–266, 2023, doi: 10.1007/978-3-031-40292-0\_21.
- [21] R. Taniguchi, R. Hoshino, dan Y. Kano, "Legal Question Answering System Using FrameNet," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11717 LNAI, hlm. 193–206, 2019, doi: 10.1007/978-3-030-31605-1\_15.
- [22] W. Huang *dkk.*, "Generating Reasonable Legal Text through the Combination of Language Modeling and Question Answering," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 4, hlm. 3687–3693, Jul 2020, doi: 10.24963/IJCAI.2020/510.
- [23] A. Abdallah, B. Piryani, dan A. Jatowt, "Exploring the state of the art in legal QA systems," *J Big Data*, vol. 10, no. 1, hlm. 1–33, Des 2023, doi: 10.1186/S40537-023-00802-8/TABLES/4.
- [24] G. Castrillo *dkk.*, "Questions and Answers on Legal Texts Based on BERT-BiGRU," *J Phys Conf Ser*, vol. 1828, no. 1, hlm. 012035, Feb 2021, doi: 10.1088/1742-6596/1828/1/012035.
- [25] R. F. Noviyani dan N. A. Rakhmawati, "rochimfn/dataset-qa: Release v1.0.1", doi: 10.5281/ZENODO.6818393.