



DAFTAR PUSTAKA

- Ahmed Awan, M.N., Khan, S., Latif, K. dan Khattak, A.M., 2019, A New Approach to Information Extraction in User-Centric E-Recruitment Systems, *Applied Sciences*, [Online] 9 (14), 2852, tersedia di DOI:10.3390/app9142852.
- Alshammari, N. dan Alanazi, S., 2020, An Arabic dataset for disease named entity recognition with multi-annotation schemes, *Data*, [Online] 5 (3), 1–8, tersedia di DOI:10.3390/data5030060.
- Bank, W., 2020, *The Promise of Education in Indonesia*, [Online]. tersedia di DOI:10.1596/34807.
- Barducci, A., Iannaccone, S., La Gatta, V., Moscato, V., Sperli, G. dan Zavota, S., 2022, An End-to-End Framework for Information Extraction from Italian Resumes, *Expert Systems with Applications*, [Online] 210 (August), 118487, tersedia di DOI:10.1016/j.eswa.2022.118487.
- Bennett, A.F., Escudero, V., Liepmann, H. dan Podjanin, A., 2022, *Using Online Vacancy and Job Applicants ' Data to Study Skills Dynamics*, International Labour Organization.
- Bhola, A., Halder, K., Prasad, A. dan Kan, M.-Y., 2021, Retrieving Skills from Job Descriptions: A Language Model Based Extreme Multi-label Classification Framework, *Proceedings of the 28th International Conference on Computational Linguistics*, [Online], 2021 hal. 5832–5842, tersedia di DOI:10.18653/v1/2020.coling-main.513.
- Bhowmick, R.S., Ganguli, I. dan Sil, J., 2020, Improvised Transformer Network for NER on Low Resource English-Hindi Code-Mixed Language from Scratch, *2020 7th Intl. Conference on Soft Computing & Machine Intelligence*, 2020 hal. 265–269,
- Boselli, R. dan dkk, 2018, Classifying online Job Advertisements through Machine Learning, *Future Generation Computer Systems*, 86319–328,
- Cao, L. dan Zhang, J., 2021, Skill Requirements Analysis for Data Analysts Based on Named Entities Recognition, *2nd International Conference on Big Data and Informatization Education (ICBDIE)*, [Online], 2021 hal. 64–68, tersedia di DOI:10.1109/ICBDIE52740.2021.00023.
- Chernova, M., 2020, Occupational skills extraction with FinBERT, *Tesis*, Metropolia University of Applied Sciences.
- Dai, Z., Yang, Z., Yang, Y., Carbonell, J., Le, Q. V. dan Salakhutdinov, R., 2019, Transformer-XL: Attentive language models beyond a fixed-length context, *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, [Online], 2019 hal. 2978–2988, tersedia di DOI:10.18653/v1/p19-1285.
- Debbarma, A., Bhattacharya, P. dan Purkayastha, B.S., 2019, Named entity



recognition for a low resource language, *International Journal of Recent Technology and Engineering*, [Online] 8 (3), 587–590, tersedia di DOI:10.35940/ijrte.B2085.098319.

Decorte, J.-J., Van Hautte, J., Demeester, T. dan Develder, C., 2021, JobBERT: Understanding Job Titles through Skills, *The International Workshop on Fair, Effective And Sustainable Talent management using data science*, [Online], 2021 hal. 1–9, tersedia di <http://arxiv.org/abs/2109.09605>.

Deng, L. dan Liu, Y., 2018, *Deep Learning in Conversational Language Understanding*, Springer Nature Singapore., [Online]. tersedia di DOI:10.1007/978-981-10-5209-5.

Devlin, J., Chang, M.W., Lee, K. dan Toutanova, K., 2019, BERT: Pre-training of deep bidirectional transformers for language understanding, *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 1 (Mlm), 4171–4186,

Erwantiningsih, E. dan Hartatik, 2021, Kinerja Ditinjau dari Keterampilan dan Pengalaman Kerja, *Meningkatkan Kinerja Manajemen Sumber Daya Manusia Konsep dan Studi Kasus*, hal. 89–109,

Fareri, S., Melluso, N., Chiarello, F. dan Fantoni, G., 2021, SkillNER : Mining and Mapping Soft Skills from Any Text, *Expert Systems With Applications*, [Online] 184 (July), 115544, tersedia di DOI:10.1016/j.eswa.2021.115544.

Gugnani, A. dan Misra, H., 2020, Implicit Skills Extraction Using Document Embedding and Its Use in Job Recommendation, *The Thirty-Second Innovative Applications of Artificial Intelligence Conference (IAAI-20)*, 2020 hal. 13286–13293,

Gupta, T., Zaki, M. dan Krishnan, N.M.A., 2022, MatSciBERT : A Materials Domain Language Model for Text Mining and Information Extraction, *Computational Materials*, [Online] 8 (102), 1–11, tersedia di DOI:10.1038/s41524-022-00784-w.

Hovy, E. dan Lavid, J., 2010, Towards a ‘science’ of corpus annotation: a new methodological challenge for corpus linguistics, *International journal of translation*, 22 (1), 13–36,

Huang, Z., Xu, P., Liang, D. dan Xiang, B., 2020, Improve Transformer Models with Better Relative Position Embeddings, *Findings of the Association for Computational Linguistics: EMNLP 2020*, [Online], 2020 hal. 3327–3335, tersedia di DOI:10.18653/v1/2020.findings-emnlp.298.

ILO, 2015, *Indonesia : Sistem informasi dan layanan*.

ILO, 2017, *Indonesia 2017*, International Labour Organization.

ILO, 2020, *World Employment and Social Outlook: Trends 2020*, [Online]. tersedia di https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_734455.pdf.



- ISCO-08, 2008, International Standard Classification of Occupations, *Department of Statistics International Labour Office*, [Online] I, tersedia di <http://www.ilo.org/public/english/bureau/stat/isco/index.htm>.
- Jati, B.S., Widyawan, S. dan Muhammad Nur Rizal, S.T., 2021, Multilingual Named Entity Recognition Model for Indonesian Health Insurance Question Answering System, *3rd International Conference on Information and Communications Technology (ICOIACT)*, [Online], 2021 hal. 180–184, tersedia di DOI:10.1109/icoiact50329.2020.9332027.
- Kao, A. dan Poteet, S.R., 2007, *Natural Language Processing and Text Mining*, Springer-Verlag London Limited 2007., [Online]. tersedia di DOI:10.1007/978-3-030-23136-1_4.
- Ke, G.P.E. in L.P.-T., He, D. dan Liu, T., 2021, Rethinking Positional Encoding in Language Pre-Training, *ICLR 2021*, 2021 hal. 1–14,
- Kementerian Ketenagakerjaan dan Badan Pusat Statistik, 2014, *Klasifikasi Baku Jabatan Indonesia*, Kementerian Ketenagakerjaan dan Badan Pusat Statistik Indonesia.
- Ketenagakerjaan RI, M., 2016, Peraturan Menteri No 39 Tahun 2016 tentang Penempatan Tenaga Kerja. *Kementerian Ketenagakerjaan*. [Online]. hal.98. tersedia di https://jdih.kemnaker.go.id/asset/data_puu/Permen_39_2016.pdf.
- Kim, H. dan Kim, H., 2021, Fine-grained Named Entity Recognition Using a Multi-Stacked Feature Fusion and Dual-Stacked Output in Korean, *Applied Sciences (Switzerland)*, [Online] 11 (22), tersedia di DOI:10.3390/app112210795.
- Koto, F., Lau, J.H. dan Baldwin, T., 2021a, IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization, *Empirical Methods in Natural Language Processing (EMNLP 2021), Dominican Republic (virtual)*,
- Koto, F., Rahimi, A., Lau, J.H. dan Baldwin, T., 2021b, IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP, *Proceedings of the 28th International Conference on Computational Linguistics*, [Online], 2021 hal. 757–770, tersedia di DOI:10.18653/v1/2020.coling-main.66.
- Lauder, H. dan Mayhew, K., 2020, Higher education and the labour market : an introduction, *Oxford Review of Education*, [Online] 46 (1), 1–9, tersedia di DOI:10.1080/03054985.2019.1699714.
- Li, J., Sun, A., Han, J. dan Li, C., 2020, A Survey on Deep Learning for Named Entity Recognition, *IEEE Transactions on Knowledge and Data Engineering, 2020*, [Online] tersedia di DOI:10.1109/tkde.2020.2981314.
- Li, T., Guo, Y. dan Ju, A., 2019, A Self-Attention-Based Approach for Named Entity Recognition in Cybersecurity, *Proceedings - 2019 15th International Conference on Computational Intelligence and Security, CIS 2019*, [Online], 2019 IEEE., hal. 147–150, tersedia di DOI:10.1109/CIS.2019.00039.



- Liang, D., Nguyen, T.Q. dan Kirchhoff, K., 2020, Masked Language Model Scoring, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020 hal. 2699–2712,
- Liu, Z., Huang, D., Huang, K., Li, Z. dan Zhao, J., 2020, FinBERT : A Pre-trained Financial Language Representation Model for Financial Text Mining, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)*, 2020 hal. 4513–4519,
- Lo, K., 2019, SCIBERT : A Pretrained Language Model for Scientific Text, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 2019 hal. 3615–3620,
- Mendbayar, K. dan Aono, M., 2021, Token Level Evaluation and Feature Enhancer for Transformer-based Models, *Proceedings - 2021 8th International Conference on Advanced Informatics: Concepts, Theory, and Application, ICAICTA 2021*, [Online] 1–6, tersedia di DOI:10.1109/ICAICTA53211.2021.9640269.
- Mikolov, T., Chen, K., Corrado, G. dan Dean, J., 2013, Distributed Representations of Words and Phrases and their Compositionality, *the 26th International Conference on Neural Information Processing Systems*, 2013 hal. 1–9,
- Muzam, J., 2023, *The Challenges of Modern Economy on the Competencies of Knowledge Workers*, Springer US., [Online]. tersedia di DOI:10.1007/s13132-022-00979-y.
- Nadeau, D. dan Sekine, S., 2007, A survey of named entity recognition and classification, *International Journal of Linguistics and Language Resources*, [Online] 30 (1), 3–26, tersedia di DOI:10.1075/li.30.1.03nad.
- Nambiar, D., Karki, S., Rahardiani, D., Putri, M. dan Singh, K., 2019, Study on Skills for The Future in Indonesia, *Oxford Policy Management*, [Online] (July), 1–117, tersedia di www.opml.co.uk.
- Nielsen, M., 2015, *Neural Networks and Deep Learning*, Determination Press., [Online]. tersedia di <http://neuralnetworksanddeeplearning.com>, diakses 19 November 2019.
- Nouvel, D., Ehrmann, M. dan Rosset, S., 2016, *Named Entities for Computational Linguistic*, John Wiley & Sons, Inc.
- Nyqvist, A., 2021, Bootstrapping Annotated Job Ads using Named Entity Recognition and Swedish Language Models, *Tesis*, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology Sweden.
- Olson, M.L., Zhang, L. dan Yu, C., 2020, Adapting Pretrained Language Models For Long Document Classification, *ICLR 2020*, 2020 hal. 1–12,
- Palmer, D.D., 2010, Text Preprocessing, *Handbook of Natural Language Processing*, hal. 3–8,



- Robbins, S.P. dan Coulter, M., 2019, *Management*.
- Shaw, P., Uszkoreit, J. dan Vaswani, A., 2018, Self-attention with relative position representations, *NAACL HLT 2018 - 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, [Online], 2018 hal. 464–468, tersedia di DOI:10.18653/v1/n18-2074.
- Suhas, H. dan Manjunath, A., 2020, Differential Hiring using a Combination of NER and Word Embedding, *International Journal of Recent Technology and Engineering*, [Online] 9 (1), 1344–1349, tersedia di DOI:10.35940/ijrte.a2400.059120.
- Tim KBI, 2008, *Kamus Bahasa Indonesia*, Edisi ke-e, Pusat Bahasa, Departemen Pendidikan Nasional, Jakarta.
- Tran, H.T., Vo, H.H.P. dan Luu, S.T., 2021, Predicting Job Titles from Job Descriptions with Multi-label Text Classification, *Proceedings - 2021 8th NAFOSTED Conference on Information and Computer Science, NICS 2021*, [Online], 2021 hal. 513–518, tersedia di DOI:10.1109/NICS54270.2021.9701541.
- Vaswani, A., 2017, Attention Is All You Need, *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 2017* hal. 1–11,
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. dan Polosukhin, I., 2017, Attention is all you need, *Advances in Neural Information Processing Systems*, 2017 hal. 5999–6009,
- Wang, C., Nulty, P. dan Lillis, D., 2020, A Comparative Study on Word Embeddings in Deep Learning for Text Classification, *Proceedings of the 4th International Conference on Natural Language Processing and Information*, [Online], 2020 hal. 37–46, tersedia di DOI:10.1145/3443279.3443304.
- Wang, P., Fang, J. dan Reinspach, J., 2021, CS-BERT : a Pretrained Model for Customer Service Dialogues, *Proceedings of the 3rd Workshop on Natural Language Processing for Conversational AI*, [Online], 2021 Association for Computational Linguistics., hal. 130--142, tersedia di DOI:10.18653/v1/2021.nlp4convai-1.13.
- Wilie, B., Vincentio, K., Winata, G.I., Cahyawijaya, S., Li, X., Lim, Z.Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S. dan Purwarianti, A., 2020, IndoNLU : Benchmark and Resources for Evaluating Indonesian Natural Language Understanding, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 2020 hal. 843–857,
- Wu, C., Wu, F. dan Huang, Y., 2021, DA-Transformer: Distance-aware Transformer, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language*



- Technologies*, 2021 hal. 2059–2068,
- Wu, Y. dkk., 2016, Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, *CoRR*, [Online] 1609 (08144), 1–23, tersedia di <http://arxiv.org/abs/1609.08144>.
- Yan, B. dan Pei, M., 2022, Clinical-BERT : Vision-Language Pre-Training for Radiograph Diagnosis and Reports Generation, *The Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22)*, (Mlm),
- Yan, H., Deng, B., Li, X. dan Qiu, X., 2019, Tener: Adapting Transformer Encoder for Named Entity Recognition, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, [Online], 2019 Association for Computational Linguistics, Florence (Italy)., hal. 1–10, tersedia di DOI:ArXiv ID 1911.04474.
- Yang, Z., Dai, Z., Yang, Y. dan Carbonell, J., 2019, XLNet: Generalized Autoregressive Pretraining for Language Understanding, *33rd Conference on Neural Information Processing Systems*, 2019 hal. 1–18,
- Yu, H., Mao, X.L., Chi, Z., Wei, W. dan Huang, H., 2020, A robust and domain-adaptive approach for low-resource named entity recognition, *Proceedings - 11th IEEE International Conference on Knowledge Graph, ICKG 2020*, [Online] 297–304, tersedia di DOI:10.1109/ICBK50248.2020.00050.
- Zhang, M., Jensen, K.N., van der Goot, R. dan Plank, B., 2022a, Skill Extraction from Job Postings using Weak Supervision, *CEUR Workshop Proceedings*, 2022 hal.
- Zhang, M., Jensen, K.N., Sonniks, S.D. dan Plank, B., 2022b, SKILLSPAN: Hard and Soft Skill Extraction from English Job Postings, *NAACL 2022 - 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, [Online], 2022 hal. 4962–4984, tersedia di DOI:10.18653/v1/2022.naacl-main.366.
- Zhang, Z. dan Zhang, R., 2019, Combined Self-attention Mechanism for Biomedical Event Trigger Identification, *Proceedings - 2019 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2019*, [Online] 1009–1012, tersedia di DOI:10.1109/BIBM47256.2019.8983274.
- Zhao, M. et al, 2015, SKILL: A system for skill identification and normalization, *Proceedings of the National Conference on Artificial Intelligence*, 54012–4017,
- Zheng, Z., Lu, X.Z., Chen, K.Y., Zhou, Y.C. dan Lin, J.R., 2022, Pretrained Domain-Specific Language Model for Natural Language Processing Tasks in The AEC Domain, *Computers in Industry*, [Online] 142103733, tersedia di DOI:10.1016/j.compind.2022.103733.
- Zukov-Gregoric, A., Bachrach, Y., Minkovsky, P., Coope, S. dan Maksak, B., 2018, Neural named entity recognition using a self-attention mechanism,



UNIVERSITAS
GADJAH MADA

Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, [Online], 2018 hal. 652–656, tersedia di DOI:10.1109/ICTAI.2017.00104.