REFERENCES

[1] G. Gygax and D. Arneson, *dungeons & dragons*. Tactical Studies Rules Lake Geneva, WI, 1974, vol. 19.

[2] E. Aarseth, "From hunt the wumpus to everquest: Introduction to quest theory," in *International Conference on Entertainment Computing*. Springer, 2005, pp. 496–506.

[3] A. Stent, R. Prasad, and M. Walker, "Trainable sentence planning for complex information presentations in spoken dialog systems," in *Proceedings of the 42nd Annual Meeting of the ACL)*, 2004, pp. 79–86.

[4] S. Busemann and H. Horacek, "A flexible shallow approach to text generation," *arXiv preprint cs/9812018*, 1998.

[5] E. Reiter and R. Dale, "Building applied natural language generation systems," *Natural Language Engineering*, vol. 3, no. 1, pp. 57–87, 1997.

[6] S. W. McRoy, S. Channarukul, and S. S. Ali, "An augmented template-based approach to text realization," *Natural Language Engineering*, vol. 9, no. 4, pp. 381–420, 2003.

[7] R. Dale, S. Geldof, and J.-P. Prost, "Coral: Using natural language generation for navigational assistance," in *Proceedings of the 26th Australasian computer science conference-Volume 16*, 2003, pp. 35–44.

[8] R. Turner, S. Sripada, and E. Reiter, "Generating approximate geographic descriptions," in *Empirical methods in natural language generation*. Springer, 2009, pp. 121–140.

[9] E. Reiter, S. Sripada, J. Hunter *et al.*, "Choosing words in computer-generated weather forecasts," *AI*, vol. 167, no. 1-2, pp. 137–169, 2005.

[10] R. Barzilay and M. Lapata, "Collective content selection for concept-to-text generation," in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, 2005, pp. 331–338.

[11] P. Liang, M. I. Jordan, and D. Klein, "Learning semantic correspondences with less supervision," in *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP*, 2009, pp. 91–99.

[12] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever *et al.*, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.

[13] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.

[14] S. Sato, "A challenge to the third hoshi shinichi award," in *Proceedings of the INLG 2016 Workshop on Computational Creativity in Natural Language Generation*. Edinburgh, UK: Association for Computational Linguistics, Sep. 2016, pp. 31–35. [Online]. Available: https://aclanthology.org/W16-5505

[15] ——, "Haoribricks3: A domain-specific language for japanese sentence composition," *Journal of Natural Language Processing*, vol. 27, no. 2, pp. 411–444, 2020.

[16] I. M. Laclaustra, J. Ledesma, G. Méndez, and P. Gervás, "Kill the dragon and rescue the princess: Designing a plan-based multi-agent story generator." in *ICCC*, 2014, pp. 347–350.

[17] P. Gervás, "Ucm submission to the surface realization challenge," in *Proceedings of the 13th European workshop on natural language generation*, 2011, pp. 239–241.

[18] W. M. Si, P. Ammanabrolu, and M. O. Riedl, "Telling stories through multi-user dialogue by modeling character relations," *arXiv preprint arXiv:2105.15054*, 2021.

[19] R. Rameshkumar and P. Bailey, "Storytelling with dialogue: A Critical Role Dungeons and Dragons Dataset," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, Jul. 2020, pp. 5121–5134. [Online]. Available: https://aclanthology.org/2020.acl-main.459

[20] S. Roller, E. Dinan, N. Goyal, D. Ju, M. Williamson, Y. Liu, J. Xu, M. Ott, K. Shuster, E. M. Smith *et al.*, "Recipes for building an open-domain chatbot," *arXiv preprint arXiv:2004.13637*, 2020.

[21] N. Mostafazadeh, N. Chambers, X. He, D. Parikh, D. Batra, L. Vanderwende, P. Kohli, and J. Allen, "A corpus and evaluation framework for deeper understanding of commonsense stories," *arXiv preprint arXiv:1604.01696*, 2016.

[22] T. Liu, K. Wang, L. Sha, B. Chang, and Z. Sui, "Table-to-text generation by structure-aware seq2seq learning," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[23] H. Chen, H. Takamura, and H. Nakayama, "SciXGen: A scientific paper dataset for context-aware text generation," in *Findings of the Association for Computational Linguistics: EMNLP 2021*. Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 1483–1492. [Online]. Available: https://aclanthology.org/2021.findings-emnlp.128

[24] H. Zhang, H. Song, S. Li, M. Zhou, and D. Song, "A survey of controllable text generation using transformer-based pre-trained language models," *arXiv preprint arXiv:2201.05337*, 2022.

[25] A. Amin-Nejad, J. Ive, and S. Velupillai, "Exploring transformer text generation for medical dataset augmentation," in *Proceedings of the Twelfth Language Resources and Evaluation Conference*. Marseille, France: European Language Resources Association, May 2020, pp. 4699–4708. [Online]. Available: https://aclanthology.org/2020.lrec-1.578

[26] J. Guan, R. Li, S. Yu, and X. Zhang, "Generation of synthetic electronic medical record text," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2018, pp. 374–380.

[27] D. Hüske-Kraus, "Text generation in clinical medicine–a review," *Methods of information in medicine*, vol. 42, no. 01, pp. 51–60, 2003.

[28] R. Lebret, D. Grangier, and M. Auli, "Neural text generation from structured data with application to the biography domain," *arXiv preprint arXiv:1603.07771*, 2016.

[29] V. Munigala, A. Mishra, S. G. Tamilselvam, S. Khare, R. Dasgupta, and A. Sankaran, "Persuaide! an adaptive persuasive text generation system for fashion domain," in *Companion Proceedings of the The Web Conference 2018*, 2018, pp. 335–342.

[30] X. Zhang, Y. Jiang, Y. Shang *et al.*, "DSGPT: Domain-Specific Generative Pre-Training of Transformers for Text Generation in E-commerce Title and Review Summarization," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 2146–2150.

[31] I. Rajapaksha, C. R. Mudalige *et al.*, "Rule-Based Approach for Party-Based Sentiment Analysis in Legal Opinion Texts," in *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2020, pp. 284–285.

[32] K. Srinivasan, K. Raman *et al.*, "Wit: Wikipedia-based image text dataset for multimodal multilingual machine learning," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 2443–2449.

[33] D. G. Ghalandari, C. Hokamp *et al.*, "A large-scale multi-document summarization dataset from the wikipedia current events portal," *arXiv preprint arXiv:2005.10070*, 2020.

[34] P. Rajpurkar, J. Zhang *et al.*, "Squad: 100,000+ questions for machine comprehension of text," *arXiv preprint arXiv:1606.05250*, 2016.

[35] J. Kreutzer, I. Caswell *et al.*, "Quality at a glance: An audit of web-crawled multilingual datasets," *TACL*, vol. 10, pp. 50–72, 2022.

[36] J. Togelius, G. N. Yannakakis, K. O. Stanley, and C. Browne, "Search-based procedural content generation: A taxonomy and survey," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 3, no. 3, pp. 172–186, 2011.

[37] A. M. Smith and M. Mateas, "Answer set programming for procedural content generation: A design space approach," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 3, no. 3, pp. 187–200, 2011.

[38] B. M. Viana and S. R. dos Santos, "A survey of procedural dungeon generation," in *2019 18th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. IEEE, 2019, pp. 29–38.

[39] C. Ashmore and M. Nitsche, "The quest in a generated world." in *DiGRA conference*. Citeseer, 2007.

[40] M. Hendrikx, S. Meijer, J. V. D. Velden, and A. Iosup, "Procedural content generation for games," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 9, no. 1, p. 1–22, 2013.

[41] G. Smith, "An analog history of procedural content generation." in *FDG*, 2015.

[42] R. Khaled, M. J. Nelson, and P. Barr, "Design metaphors for procedural content generation in games," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. Association for Computing Machinery, 2013, p. 1509–1518.

[43] N. Shaker, A. Liapis, J. Togelius, R. Lopes, and R. Bidarra, "Constructive generation methods for dungeons and levels," *Procedural Content Generation in Games Computational Synthesis and Creative Systems*, p. 31–55, 2016.

[44] M. Niemann and M. Preuß, "Constructive generation methods for dungeons," in *Seminar-Thesis in Procedural Content Generation for Games. Westfalische Wilhelms Universitat Munster*, 2015.

[45] R. Pereira de Araujo and V. T. Souto, "Game worlds and creativity: The challenges of procedural content generation," *Lecture Notes in Computer Science*, p. 443–455, 2017.

[46] E. Olsson and E. Grevillius, "Mixed-initiative quest generation," 2020.

[47] A. Baldwin, S. Dahlskog, J. M. Font, and J. Holmberg, "Towards pattern-based mixed-initiative dungeon generation," in *Proceedings of the 12th International Conference on the Foundations of Digital Games*, 2017, pp. 1–10.

[48] A. Alvarez and J. Font, "Learning the designer's preferences to drive evolution," *arXiv preprint arXiv:2003.03268*, 2020.

[49] Y.-S. Lee and S.-B. Cho, "Context-aware petri net for dynamic procedural content generation in role-playing game," *IEEE Computational Intelligence Magazine*, vol. 6, no. 2, pp. 16–25, 2011.

[50] C. Brom and A. Abonyi, "Petri-nets for game plot," in *Proceedings of AISB artificial intelligence and simulation behaviour convention, Bristol*, vol. 3, 2006, pp. 6–13.

[51] W. M. van der Aalst, "Process discovery: Capturing the invisible," *IEEE Computational Intelligence Magazine*, vol. 5, no. 1, pp. 28–41, 2010.

[52] J. L. Peterson, *Petri net theory and the modeling of systems*. Prentice Hall PTR, 1981.

[53] J. Doran and I. Parberry, "A prototype quest generator based on a structural analysis of quests from four mmorpgs," in *PCGames '11*, 2011.

[54] N. Lazzaro, "Why we play games: Four keys to more emotion in player experiences," in *Proceedings of GDC*, vol. 306, 2004, pp. 1–8.

[55] E. Horvitz, "Mixed-initiative interaction," *IEEE Intelligent Systems*, pp. 14–24, September 1999. [Online]. Available: https://www.microsoft.com/en-us/research/publication/mixed-initiative-interaction/

[56] J. E. Allen, C. I. Guinn, and E. Horvtz, "Mixed-initiative interaction," *IEEE Intelligent Systems and their Applications*, vol. 14, no. 5, pp. 14–23, 1999.

[57] A. Baldwin, S. Dahlskog, J. M. Font, and J. Holmberg, "Mixed-initiative procedural generation of dungeons using game design patterns," in *2017 IEEE conference on computational intelligence and games (CIG)*. IEEE, 2017, pp. 25–32.

[58] A. Alvarez, S. Dahlskog, J. Font, J. Holmberg, and S. Johansson, "Assessing aesthetic criteria in the evolutionary dungeon designer," in *Proceedings of the 13th International Conference on the Foundations of Digital Games*, 2018, pp. 1–4.

[59] S. Dahlskog and J. Togelius, "A multi-level level generator," in *2014 IEEE Conference on Computational Intelligence and Games*. IEEE, 2014, pp. 1–8.

[60] A. Alvarez, S. Dahlskog, J. Font, and J. Togelius, "Empowering quality diversity in dungeon design with interactive constrained map-elites," in *2019 IEEE Conference on Games (CoG)*. IEEE, 2019, pp. 1–8.

[61] A. Liapis, G. N. Yannakakis, and J. Togelius, "Designer modeling for sentient sketchbook," in *2014 IEEE Conference on Computational Intelligence and Games*. IEEE, 2014, pp. 1–8.

[62] G. N. Yannakakis, A. Liapis, and C. Alexopoulos, "Mixed-initiative cocreativity," in *International Conference on Foundations of Digital Games*, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:7551405

[63] A. Alvarez, J. M. M. F. Fernandez, S. Dahlskog, and J. Togelius, "Interactive constrained map-elites: Analysis and evaluation of the expressiveness of the feature dimensions," *IEEE Transactions on Games*, 2020.

[64] J. Dormans, "Adventures in level design: generating missions and spaces for action adventure games," in *Proceedings of the 2010 workshop on procedural content generation in games*, 2010, pp. 1–8.

[65] D. Ashlock and C. McGuinness, "Automatic generation of fantasy role-playing modules," in *2014 IEEE Conference on Computational Intelligence and Games*. IEEE, 2014, pp. 1–8.

[66] A. Sullivan, M. Mateas, and N. Wardrip-Fruin, "Questbrowser: Making quests playable with computer-assisted design," 2009.

[67] A. Rollings and D. Morris, *Game Architecture and Design*, ser. NRG Series. New Riders, 2004. [Online]. Available: https://books.google.lk/books?id=j45QAAAAMAAJ

[68] J. Howard, *Quests: Design, theory, and history in games and narratives*. AK Peters/CRC Press, 2022.

[69] M. Nitsche, C. Ashmore, W. Hankinson, R. Fitzpatrick, J. Kelly, and K. Margenau, "Designing procedural game spaces: A case study," *Proceedings of FuturePlay*, vol. 2006, 2006.

[70] M. Brenner, "Creating dynamic story plots with continual multiagent planning," in *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.

[71] E. Soares de Lima, B. Feijó, and A. L. Furtado, "Hierarchical generation of dynamic and nondeterministic quests in games," in *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*, 2014, pp. 1–10.

[72] Y. Huang, Y. Zhang, O. Elachqar, and Y. Cheng, "Inset: Sentence infilling with inter-sentential transformer," *arXiv preprint arXiv:1911.03892*, 2019.

[73] W. Zhu, Z. Hu, and E. Xing, "Text infilling," *arXiv preprint arXiv:1901.00158*, 2019.

[74] C. Donahue, M. Lee, and P. Liang, "Enabling language models to fill in the blanks," *arXiv preprint arXiv:2005.05339*, 2020.

[75] W. L. Taylor, ""cloze procedure": A new tool for measuring readability," *Journalism quarterly*, vol. 30, no. 4, pp. 415–433, 1953.

[76] D. Ippolito, D. Grangier, C. Callison-Burch, and D. Eck, "Unsupervised hierarchical story infilling," in *Proceedings of the First Workshop on Narrative Understanding*, 2019, pp. 37–43.

[77] Y. Xiao, L. Liu, G. Huang, Q. Cui, S. Huang, S. Shi, and J. Chen, "BiTIIMT: A bilingual text-infilling method for interactive machine translation," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 1958–1969. [Online]. Available: https://aclanthology.org/2022.acl-long.138

[78] Y.-S. Shih, W.-C. Chang, and Y. Yang, "Xl-editor: Post-editing sentences with xlnet," *arXiv preprint arXiv:1910.10479*, 2019.

[79] Y. Assael, T. Sommerschield, and J. Prag, "Restoring ancient text using deep learning: a case study on greek epigraphy," *arXiv preprint arXiv:1910.06262*, 2019.

[80] X. Qiu, T. Sun, Y. Xu, Y. Shao, N. Dai, and X. Huang, "Pre-trained models for natural language processing: A survey," *Science China Technological Sciences*, vol. 63, no. 10, pp. 1872–1897, 2020.

[81] J. Ou, N. Weir, A. Belyy, F. Yu, and B. Van Durme, "Infillmore: Frame-guided language generation with bidirectional context," *arXiv preprint arXiv:2103.04941*, 2021.

[82] J. Li, R. Jia, H. He, and P. Liang, "Delete, retrieve, generate: a simple approach to sentiment and style transfer," *arXiv preprint arXiv:1804.06437*, 2018.

[83] C. Shaoul, R. H. Baayen, and C. F. Westbury, "N-gram probability effects in a cloze task," *The mental lexicon*, vol. 9, no. 3, pp. 437–472, 2014.

[84] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *Advances in neural information processing systems*, vol. 27, 2014.

[85] X. Wu, T. Zhang, L. Zang, J. Han, and S. Hu, ""' mask and infill": Applying masked language model to sentiment transfer," *arXiv preprint arXiv:1908.08039*, 2019.

[86] D. Fried, A. Aghajanyan, J. Lin, S. Wang, E. Wallace, F. Shi, R. Zhong, W.-t. Yih, L. Zettlemoyer, and M. Lewis, "Incoder: A generative model for code infilling and synthesis," *arXiv preprint arXiv:2204.05999*, 2022.

[87] A. Griffin, "Ai deemed "too dangerous to release" makes it out into the world," Nov 2019. [Online]. Available: https://www.independent.co.uk/tech/ai-artificial-intelligence-dangerous-text-gpt2-elon-musk-a9192121.html

[88] OpenAI, "Chatgpt," https://openai.com/gpt, 2022, accessed on 2024-03-01.

[89] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[90] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[91] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423

[92] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever *et al.*, "Improving language understanding by generative pre-training," 2018.

[93] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat *et al.*, "Gpt-4 technical report," *arXiv preprint arXiv:2303.08774*, 2023.

[94] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. d. l. Casas, F. Bressand, G. Lengyel, G. Lample, L. Saulnier *et al.*, "Mistral 7b," *arXiv preprint arXiv:2310.06825*, 2023.

[95] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar *et al.*, "Llama: Open and efficient foundation language models," *arXiv preprint arXiv:2302.13971*, 2023.

[96] G. Penedo, Q. Malartic, D. Hesslow, R. Cojocaru, A. Cappelli, H. Alobeidli, B. Pannier, E. Almazrouei, and J. Launay, "The refinedweb dataset for falcon llm: outperforming curated corpora with web data, and web data only," *arXiv preprint arXiv:2306.01116*, 2023.

[97] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P. J. Liu *et al.*, "Exploring the limits of transfer learning with a unified text-to-text transformer." *J. Mach. Learn. Res.*, vol. 21, no. 140, pp. 1–67, 2020.

[98] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, Jul. 2020, pp. 7871–7880. [Online]. Available: https://aclanthology.org/2020.acl-main.703

[99] Z. Feng, D. Guo, D. Tang, N. Duan, X. Feng, M. Gong, L. Shou, B. Qin, T. Liu, D. Jiang *et al.*, "Codebert: A pre-trained model for programming and natural languages," *arXiv preprint arXiv:2002.08155*, 2020.

[100] A. Aghajanyan, B. Huang, C. Ross, V. Karpukhin, H. Xu, N. Goyal, D. Okhonko, M. Joshi, G. Ghosh, M. Lewis *et al.*, "Cm3: A causal masked multimodal model of the internet," *arXiv preprint arXiv:2201.07520*, 2022.

[101] Q. Wu, G. Bansal, J. Zhang, Y. Wu, S. Zhang, E. Zhu, B. Li, L. Jiang, X. Zhang, and C. Wang, "Autogen: Enabling next-gen llm applications via multi-agent conversation framework," *arXiv preprint arXiv:2308.08155*, 2023.

[102] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel *et al.*, "Retrieval-augmented generation for knowledge-intensive nlp tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[103] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "Lora: Low-rank adaptation of large language models," *arXiv preprint arXiv:2106.09685*, 2021.

[104] T. Dettmers, "8-bit approximations for parallelism in deep learning," *arXiv preprint arXiv:1511.04561*, 2015.

[105] T. Dettmers and L. Zettlemoyer, "The case for 4-bit precision: k-bit inference scaling laws," in *International Conference on Machine Learning*. PMLR, 2023, pp. 7750–7774.

[106] J. Chee, Y. Cai, V. Kuleshov, and C. M. De Sa, "Quip: 2-bit quantization of large language models with guarantees," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[107] S. Ma, H. Wang, L. Ma, L. Wang, W. Wang, S. Huang, L. Dong, R. Wang, J. Xue, and F. Wei, "The era of 1-bit llms: All large language models are in 1.58 bits," *arXiv preprint arXiv:2402.17764*, 2024.

[108] M. Xu, W. Yin, D. Cai, R. Yi, D. Xu, Q. Wang, B. Wu, Y. Zhao, C. Yang, S. Wang *et al.*, "A survey of resource-efficient llm and multimodal foundation models," *arXiv preprint arXiv:2401.08092*, 2024.

[109] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, "Qlora: Efficient finetuning of quantized llms," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[110] T. Dettmers, M. Lewis, Y. Belkada, and L. Zettlemoyer, "Gpt3. int8 (): 8-bit matrix multiplication for transformers at scale," *Advances in Neural Information Processing Systems*, vol. 35, pp. 30 318–30 332, 2022.

[111] E. Frantar, S. Ashkboos, T. Hoefler, and D. Alistarh, "Gptq: Accurate post-training quantization for generative pre-trained transformers," *arXiv preprint arXiv:2210.17323*, 2022.

[112] G. Gerganov, "Ggml/docs/gguf.md at master · ggerganov/ggml," Nov 2023. [Online]. Available: https://github.com/ggerganov/ggml/blob/master/docs/gguf.md

[113] A. Ferrari, B. Donati, and S. Gnesi, "Detecting domain-specific ambiguities: an nlp approach based on wikipedia crawling and word embeddings," in *2017 IEEE 25th International Requirements Engineering Conference Workshops (REW)*. IEEE, 2017, pp. 393–399.

[114] V. Cozza, M. Petrocchi, and A. Spognardi, "A matter of words: NLP for quality evaluation of Wikipedia medical articles," in *International Conference on Web Engineering*. Springer, 2016, pp. 448–456.

[115] O. Ferschke, "The quality of content in open online collaboration platforms: Approaches to nlp-supported information quality management in wikipedia," Ph.D. dissertation, Technische Universität Darmstadt, 07 2014.

[116] S. P. Ponzetto and M. Strube, "Knowledge derived from wikipedia for computing semantic relatedness," *Journal of Artificial Intelligence Research*, vol. 30, pp. 181–212, 2007.

[117] T. Zesch and I. Gurevych, "Analysis of the wikipedia category graph for nlp applications," in *Proceedings of the Second Workshop on TextGraphs: Graph-Based Algorithms for Natural Language Processing*, 2007, pp. 1–8.

[118] T. Zesch, C. Müller, and I. Gurevych, "Extracting lexical semantic knowledge from wikipedia and wiktionary," in *Proceedings of the Sixth LREC*, 2008.

[119] O. Ferschke, J. Daxenberger, and I. Gurevych, "A survey of nlp methods and resources for analyzing the collaborative writing process in wikipedia," in *The People's Web Meets NLP*. Springer, 2013, pp. 121–160.

[120] V. Nastase and M. Strube, "Transforming wikipedia into a large scale multilingual concept network," *AI*, vol. 194, pp. 62–85, 2013.

[121] D. Lange, C. Böhm, and F. Naumann, "Extracting structured information from wikipedia articles to populate infoboxes," in *Proceedings of the 19th ACM international conference on Information and knowledge management*, 2010, pp. 1661–1664.

[122] D. Lamprecht, D. Dimitrov, D. Helic, and M. Strohmaier, "Evaluating and improving navigability of wikipedia: a comparative study of eight language editions," in *Proceedings of the 12th international symposium on open collaboration*, 2016, pp. 1–10.

[123] A. Peiris and N. de Silva, "Synthesis and evaluation of a domain-specific large data set for dungeons & dragons," in *Proceedings of the 36th Pacific Asia Conference on Language, Information and Computation*. Manila, Philippines: Association for Computational Linguistics, Oct. 2022, pp. 415–424. [Online]. Available: https://aclanthology.org/2022.paclic-1.46

[124] S. Whitten, "Dungeons & Dragons had its biggest year ever as Covid forced the game off tables and onto the web,"

Mar. 2021. [Online]. Available: https://www.cnbc.com/2021/03/13/dungeons-dragons-had-its-biggest-year-despite-the-coronavirus.html

[125] J. Crawford, J. Wyatt, R. J. Schwalb, and B. R. Cordell, *Player's handbook*. Wizards of the Coast LLC, 2014.

[126] O. A. Collective, "axolotl," 9 2023. [Online]. Available: https://github.com/OpenAccess-AI-Collective/axolotl

[127] M. Ibrahim, C. M. Danforth, and P. S. Dodds, "Connecting every bit of knowledge: The structure of wikipedia's first link network," *Journal of Computational Science*, vol. 19, pp. 21–30, 2017.

[128] M. Nickel and D. Kiela, "Poincaré embeddings for learning hierarchical representations," in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio *et al.*, Eds. Curran Associates, Inc., 2017, pp. 6341–6350. [Online]. Available: http://papers.nips.cc/paper/7213-poincare-embeddings-for-learning-hierarchical-representations.pdf

[129] Z. Wu and M. Palmer, "Verb semantics and lexical selection," *arXiv preprint cmp-lg/9406033*, 1994.

[130] J. J. Jiang and D. W. Conrath, "Semantic similarity based on corpus statistics and lexical taxonomy," *arXiv preprint cmp-lg/9709008*, 1997.

[131] L. Xing and M. J. Paul, "Incorporating metadata into content-based user embeddings," in *Proceedings of the 3rd Workshop on Noisy User-generated Text*. Copenhagen, Denmark: Association for Computational Linguistics, Sep. 2017, pp. 45–49. [Online]. Available: https://aclanthology.org/W17-4406

[132] H. Zhou, L. Chen, F. Shi, and D. Huang, "Learning bilingual sentiment word embeddings for cross-language sentiment classification," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Beijing, China: Association for Computational Linguistics, Jul. 2015, pp. 430–440. [Online]. Available: https://aclanthology.org/P15-1042

[133] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[134] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *International conference on machine learning*. PMLR, 2014, pp. 1188–1196.

[135] L. Gao, S. Biderman, S. Black, L. Golding, T. Hoppe, C. Foster, J. Phang, H. He, A. Thite, N. Nabeshima *et al.*, "The pile: An 800gb dataset of diverse text for language modeling," *arXiv preprint arXiv:2101.00027*, 2020.

[136] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry *et al.*, "Language models are few-shot learners," *Advances in Neural Information Processing Systems*, vol. 2020-December, 5 2020. [Online]. Available: https://arxiv.org/abs/2005.14165v4

[137] A. Peiris and N. de Silva, "Shade: Semantic hypernym annotator for domain-specific entities-dungeons and dragons domain use case," in *2023 IEEE 17th International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 2023, pp. 1–6.

[138] M. Liberman, "Language Log: Dwarves vs. Dwarfs," https://bit.ly/3JAgVpA, January 2004, (Accessed on 03/14/2023).

[139] R. Artstein, "Inter-annotator agreement," *Handbook of linguistic annotation*, pp. 297–313, 2017.

[140] R. Artstein and M. Poesio, "Inter-coder agreement for computational linguistics," *CL*, vol. 34, no. 4, pp. 555–596, 2008.

[141] E. Ouyang, Y. Li *et al.*, "Exploring n-gram character presentation in bidirectional rnn-crf for chinese clinical named entity recognition," in *CEUR workshop proceedings*, vol. 1976, 2017, pp. 37–42.

[142] I. Wijesiri, M. Gallage *et al.*, "Building a wordnet for sinhala," in *Proceedings of the seventh global wordnet conference*, 2014, pp. 100–108.

[143] F. Wu and D. S. Weld, "Automatically Refining the Wikipedia Infobox Ontology," in *WWW*, 2008, pp. 635–644.

[144] A. karpathy, "nanoGPT," 1 2023. [Online]. Available: https://github.com/karpathy/nanoGPT/

[145] Y. Wijeratne and I. Marikar, "Better question-answering models on a budget," *arXiv preprint arXiv:2304.12370*, 2023.

[146] P. Bo, "Blinkdl/rwkv-lm: 0.01," Aug. 2021. [Online]. Available: https://doi.org/10.5281/zenodo.5196577

[147] H. Chase, "LangChain," Oct. 2022. [Online]. Available: https://github.com/langchain-ai/langchain

[148] J. Moura, "crewAI," Nov. 2023. [Online]. Available: https://github.com/joaomdmoura/crewAI

[149] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and psychological measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[150] J. L. Fleiss, "Measuring nominal scale agreement among many raters." *Psychological bulletin*, vol. 76, no. 5, p. 378, 1971.

[151] K. Krippendorff, *Content analysis: An introduction to its methodology*. Sage publications, 2018.

[152] W. A. Scott, "Reliability of content analysis: The case of nominal scale coding," *Public opinion quarterly*, pp. 321–325, 1955.

[153] S. Black, G. Leo, P. Wang, C. Leahy, and S. Biderman, "GPT-Neo: Large Scale Autoregressive Language Modeling with Mesh-Tensorflow," Mar. 2021. [Online]. Available: https://doi.org/10.5281/zenodo.5297715

[154] R. Taori, I. Gulrajani, T. Zhang, Y. Dubois, X. Li, C. Guestrin, P. Liang, and T. B. Hashimoto, "Stanford alpaca: An instruction-following llama model," https://github.com/tatsu-lab/stanford_alpaca, 2023.

[155] R. Snow, B. O'Connor, D. Jurafsky, and A. Ng, "Cheap and fast – but is it good? evaluating non-expert annotations for natural language tasks," in *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. Honolulu, Hawaii: Association for Computational Linguistics, Oct. 2008, pp. 254–263. [Online]. Available: https://aclanthology.org/D08-1027