Universiteit
Leiden
The Netherlands

Knowledge discovery from patient forums: gaining novel medical insights from patient experiences

Dirkson, A.R.

Citation

Dirkson, A. R. (2022, December 6). *Knowledge discovery from patient forums: gaining novel medical insights from patient experiences*. *SIKS Dissertation Series*. Retrieved from https://hdl.handle.net/1887/3492655

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3492655

Note: To cite this publication please use the final published version (if applicable).

REFERENCES

[1] N. K. Aaronson, S. Ahmedzai, B. Bergman, M. Bullinger, A. Cull, N. J. Duez, A. Filiberti, H. Flechtner, S. B. Fleishman, J. C. Haes, S. Kaasa, M. Klee, D. Osoba, D. Razavi, P. B. Rofe, S. Schraub, K. Sneeuw, M. Sullivan, and F. Takeda. The European organization for research and treatment of cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, 85(5):365–376, 1993. ISSN 00278874. doi: 10.1093/jnci/85.5.365.

[2] ACL rolling review. Guidelines for Answering Checklist Questions, 2021. URL `https://aclrollingreview.org/responsibleNLPresearch/`.

[3] A. Akbik, D. Blythe, and R. Vollgraf. Contextual string embeddings for sequence labeling. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1638–1649, Santa Fe, New Mexico, USA, 2018. Association for Computational Linguistics.

[4] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter, and R. Vollgraf. Flair: An easy-to-use framework for state-of-the-art nlp. In *NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, pages 54–59, 2019.

[5] A. Al Hamid, M. Ghaleb, H. Aljadhey, and Z. Aslanpour. A systematic review of qualitative research on the contributory factors leading to medicine-related problems from the perspectives of adult patients with cardiovascular diseases and diabetes mellitus. *BMJ open*, 4(9):e005992, sep 2014. ISSN 2044-6055. doi: 10.1136/bmjopen-2014-005992.

[6] E. Alsentzer, J. Murphy, W. Boag, W.-H. Weng, D. Jindi, T. Naumann, and M. McDermott. Publicly available clinical BERT embeddings. In *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, pages 72–78, Minneapolis, Minnesota, USA, 2019. Association for Computational Linguistics. doi: 10.18653/v1/W19-1909.

[7] B. Alshemali and J. Kalita. Improving the reliability of deep neural networks in NLP: A review. *Knowledge-Based Systems*, 10520, 2019. doi: 10.1016/j.knosys.2019.105210.

[8] M. Alzantot, Y. Sharma, A. Elgohary, B.-J. Ho, M. B. Srivastava, and K.-W. Chang. Generating Natural Language Adversarial Examples. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2890–2896. Association for Computational Linguistics, 2018.

[9] C. Anderson, J. Krska, E. Murphy, and A. Avery. The importance of direct patient reporting of suspected adverse drug reactions: A patient perspective. *British Journal of Clinical Pharmacology*, 72(5):806–822, 2011. ISSN 03065251. doi: 10.1111/j.1365-2125.2011.03990.x.

[10] M. Anderson and A. Smith. Social Media Use in 2021. Technical Report April, Pew Research Center, 2021. URL `https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/`.

[11] V. Araujo, A. Carvallo, and D. Parra. Adversarial evaluation of bert for biomedical named entity recognition. In *Proceedings of the The Fourth Widening Natural Language Processing Workshop*, pages 79–82, 2020.

[12] T. M. Atkinson, S. J. Ryan, A. V. Bennett, A. M. Stover, R. M. Saracino, L. J. Rogak, S. T. Jewell, K. Matsoukas, Y. Li, and E. Basch. The association between clinician-based common terminology criteria for adverse events (CTCAE) and patient-reported outcomes (PRO): a systematic review. *Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer*, 24(8):3669–3676, 2016. ISSN 1433-7339. doi: 10.1007/S00520-016-3297-9.

[13] B. Audeh, F. Bellet, M. N. Beyens, A. Lillo-Le Louët, and C. Bousquet. Use of Social Media for Pharmacovigilance Activities: Key Findings and Recommendations from the Vigi4Med Project. *Drug Safety*, 43(9):835–851, 2020. ISSN 11791942. doi: 10.1007/s40264-020-00951-2.

[14] J. L. Austin. *How to do things with words*. Oxford university press, 1962.

[15] S. Aymé, A. Kole, and S. Groft. Empowerment of patients: lessons from the rare diseases community. *The Lancet*, 371(9629):2048–2051, 2008. ISSN 01406736. doi: 10.1016/S0140-6736(08)60875-2.

[16] R. Baeza-Yates. Bias on the web. *Communications of the ACM*, 2018. ISSN 00010782. doi: 10.1145/3209581.

[17] P. Bajaj, D. Campos, N. Craswell, L. Deng, J. Gao, X. Liu, R. Majumder, A. McNamara, B. Mitra, T. Nguyen, M. Rosenberg, X. Song, A. Stoica, S. Tiwary, and T. Wang. MS MARCO: A Human Generated MAchine Reading COmprehension Dataset. *ArXiv*, 2016. ISSN 16130073.

[18] S. Balasubramanian, N. Jain, G. Jindal, A. Awasthi, and S. Sarawagi. What's in a Name? Are BERT Named Entity Representations just as Good for any other Name? In *Proceedings of the 5th Workshop on Representation Learning for NLP*, pages 205–214. Association for Computational Linguistics, 2020. doi: 10.18653/v1/2020.repl4nlp-1.24.

[19] T. Baldwin, M.-C. de Marneffe, B. Han, Y.-B. Kim, A. Ritter, and W. Xu. Shared tasks of the 2015 workshop on noisy user-generated text: Twitter lexical normalization and named entity recognition. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 126–135. Association for Computational Linguistics, 2015. doi: 10.18653/v1/W15-4319.

[20] M. Basaldella, F. Liu, E. Shareghi, and N. Collier. COMETA: A corpus for medical entity linking in the social media. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3122–3137. Association for Computational Linguistics, 2020. doi: 10.18653/v1/2020.emnlp-main.253.

[21] E. Basch, X. Jia, G. Heller, A. Barz, L. Sit, M. Fruscione, M. Appawu, A. Iasonos, T. Atkinson, S. Goldfarb, A. Culkin, M. G. Kris, and D. Schrag. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *Journal of the National Cancer Institute*, 101(23):1624–1632, 2009. doi: 10.1093/JNCI/DJP386.

[22] E. Basch, B. B. Reeve, S. A. Mitchell, S. B. Clauser, L. M. Minasian, A. C. Dueck, T. R. Mendoza, J. Hay, T. M. Atkinson, A. P. Abernethy, D. W. Bruner, C. S. Cleeland, J. A. Sloan, R. Chilukuri, P. Baumgartner, A. Denicoff, D. St. Germain, A. M. O'Mara, A. Chen, J. Kelaghan, A. V. Bennett, L. Sit, L. Rogak, A. Barz, D. B. Paul, and D. Schrag. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *Journal of the National Cancer Institute*, 106(9), 2014. doi: 10.1093/JNCI/DJU244.

[23] A. Bate, R. F. Reynolds, and P. Caubel. The hope, hype and reality of Big Data for pharmacovigilance. *Therapeutic advances in drug safety*, 9(1):5–11, 2018. ISSN 2042-0986. doi: 10.1177/2042098617736422.

[24] R. Beckley. Bekli: A simple approach to twitter text normalization. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 82–86, Beijing, China, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4312.

[25] M. Beeksma, S. Verberne, A. van den Bosch, I. Hendrickx, E. Das, and S. Groenewoud. Predicting life expectancy with a recurrent neural network. *BMC Medical Informatics and Decision Making*, 19(36), 2019. doi: 10.1186/s12911-019-0775-2.

[26] Y. Belinkov and Y. Bisk. Synthetic and Natural Noise Both Break Neural Machine Translation. *ArXiv*, abs/1711.0, 2017. URL `http://arxiv.org/abs/1711.02173`.

[27] M. Belousov, W. Dixon, and G. Nenadic. Using an Ensemble of Generalised Linear and Deep Learning Models in the SMM4H 2017 Medical Concept Normalisation Task. In *Proceedings of the Second Workshop on Social Media Mining for Health Applications (SMM4H)*, 2017.

[28] I. Beltagy, K. Lo, and A. Cohan. SciBERT: A pretrained language model for scientific text. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3615–3620, Hong Kong, China, 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-1371.

[29] Y. Benjamini and Y. Hochberg. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, 5(1): 289–300, 1995.

[30] A. Benton, L. Ungar, S. Hill, S. Hennessy, J. Mao, A. Chung, C. E. Leonard, and J. H. Holmes. Identifying potential adverse effects using the web: a new approach to medical hypothesis generation HHS Public Access. *J Biomed Inform*, 44(6):989–996, 2011. doi: 10.1016/j.jbi.2011.07.005.

[31] G. Berend and E. Tasnádi. Uszeged: Correction type-sensitive normalization of english tweets using efficiently indexed n-gram statistics. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 120–125, Beijing, China, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4318.

[32] J. Berrewaerts, L. Delbecque, P. Orban, and M. Desseilles. Patient Participation and the Use of Ehealth Tools for Pharmacoviligance. *Frontiers in Pharmacology*, 7:90, 2016. ISSN 1663-9812. doi: 10.3389/fphar.2016.00090.

[33] J. Bian and F. Yu. Towards Large-scale Twitter Mining for Drug-related Adverse Events. In *SHB12*, pages 25–32, 2012. doi: 10.1145/2389707.2389713.

[34] G. Blank and C. Lutz. Representativeness of Social Media in Great Britain: Investigating Facebook, LinkedIn, Twitter, Pinterest, Google+, and Instagram. *American Behavioral Scientist*, 61(7):741–756, 2017. ISSN 15523381. doi: 10.1177/0002764217717559.

[35] C. D. Blanke, G. D. Demetri, M. von Mehren, M. C. Heinrich, B. Eisenberg, J. A. Fletcher, C. L. Corless, C. D. M. Fletcher, P. J. Roberts, D. Heinz, E. Wehre, Z. Nikolova, and H. Joensuu. Long-term results from a randomized phase II trial of standard- versus higher-dose imatinib mesylate for patients with unresectable or metastatic gastrointestinal stromal tumors expressing KIT. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 26(4):620–5, 2 2008. ISSN 1527-7755. doi: 10.1200/JCO.2007.13.4403.

[36] C. D. Blanke, C. Rankin, G. D. Demetri, C. W. Ryan, M. von Mehren, R. S. Benjamin, A. K. Raymond, V. H. C. Bramwell, L. H. Baker, R. G. Maki, M. Tanaka, J. R. Hecht, M. C. Heinrich, C. D. M. Fletcher, J. J. Crowley, and E. C. Borden. Phase III randomized, intergroup trial assessing imatinib mesylate at two dose levels in patients with unresectable or metastatic gastrointestinal stromal tumors expressing the kit receptor tyrosine kinase: S0033. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 26(4):626–32, 2008. ISSN 1527-7755. doi: 10.1200/JCO.2007.13.4452.

[37] A. Blenkinsopp, P. Wilkie, M. Wang, and P. A. Routledge. Patient reporting of suspected adverse drug reactions: A review of published literature and international experience. *British Journal of Clinical Pharmacology*, 63(2):148–156, 2007. ISSN 03065251. doi: 10.1111/j.1365-2125.2006.02746.x.

[38] T. Borkman. Experiential Knowledge: A New Concept for the Analysis of Self-Help Groups. *Social Service Review*, 50(3):445–456, 1976. ISSN 0037-7961. doi: 10.1086/643401.

[39] A. Bottomley, J. C. Reijneveld, M. Koller, H. Flechtner, K. A. Tomaszewski, E. Greimel, P. A. Ganz, J. Ringash, D. O'Connor, P. G. Kluetz, G. Tafuri, M. Grønvold, C. Snyder, C. Gotay, D. L. Fallowfield, K. Apostolidis, R. Wilson, R. Stephens, H. Schünemann, M. Calvert, B. Holzner, J. Z. Musoro, S. Wheelwright, F. Martinelli, A. C. Dueck, M. Pe, C. Coens, G. Velikova, D. Kuliś, M. J. Taphoorn, A. S. Darlington, I. Lewis, and L. van de Poll-Franse. Current state of quality of life and patient-reported outcomes research. *European journal of cancer (Oxford, England : 1990)*, 121:55–63, 2019. ISSN 1879-0852. doi: 10.1016/J.EJCA.2019.08.016.

[40] C. Bousquet, B. Dahamna, S. Guillemin-Lanne, S. J. Darmoni, C. Faviez, C. Huot, S. Katsahian, V. Leroux, S. Pereira, C. Richard, S. Schück, J. Souvignet, A. Lillo-Le Louët, and N. Texier. The Adverse Drug Reactions from Patient Reports in Social Media Project: Five Major Challenges to Overcome to Operationalize Analysis and Efficiently Support Pharmacovigilance Process. *JMIR Research Protocols*, 6(9):e179, 2017. ISSN 1929-0748. doi: 10.2196/resprot.6463.

[41] A. Brandsen, S. Verberne, M. Wansleeben, and K. Lambers. Creating a dataset for named entity recognition in the archaeology domain. In *LREC 2020 - 12th International Conference on Language Resources and Evaluation, Conference Proceedings*, pages 4573–4577, 2020. ISBN 9791095546344.

[42] S. Brosch, A. M. de Ferran, V. Newbould, D. Farkas, M. Lengsavath, and P. Tregunno. Establishing a Framework for the Use of Social Media in Pharmacovigilance in Europe. *Drug Safety*, 42(8):921–930, 2019. ISSN 11791942. doi: 10.1007/s40264-019-00811-8.

[43] A.-M. Bucur, A. Cosma, and L. P. Dinu. Sequence-to-Sequence Lexical Normalization with Multilingual Transformers. *ArXiv*, 2021. URL `https://arxiv.org/abs/2110.02869v3`.

[44] A. Bulcock, L. Hassan, S. Giles, C. Sanders, G. Nenadic, S. Campbell, and W. Dixon. Public Perspectives of Using Social Media Data to Improve Adverse Drug Reaction Reporting: A Mixed-Methods Study. *Drug Safety 2021 44:5*, 44(5):553–564, 2021. ISSN 1179-1942. doi: 10.1007/S40264-021-01042-6.

[45] M. H. Burda, M. Van Den Akker, F. Van Der Horst, P. Lemmens, and J. A. Knottnerus. Collecting and validating experiential expertise is doable but poses methodological challenges. *Journal of Clinical Epidemiology*, 72:10–15, 2016. ISSN 18785921. doi: 10.1016/j.jclinepi.2015.10.021.

[46] G. Burnage, R. Baayen, R. Piepenbrock, and H. van Rijn. *CELEX: A Guide for Users*. Centre for Lexical Information, 1990.

[47] J. Call, C. D. Walentas, J. C. Eickhoff, and N. Scherzer. Survival of gastrointestinal stromal tumor patients in the imatinib era: life raft group observational registry. *BMC Cancer*, 12(1):90, 2012. ISSN 1471-2407. doi: 10.1186/1471-2407-12-90.

[48] E. B. Carbajal-López, D. M. Juárez-García, A. Espinoza-Velazco, and G. Calderillo-Ruiz. Internet-Delivered Cognitive Behavioral Therapy and Psychoeducation

Program for Patients with Gastrointestinal Stromal Tumors. *Journal of cancer education : the official journal of the American Association for Cancer Education*, 2020. ISSN 1543-0154. doi: 10.1007/S13187-020-01866-3.

[49] J. F. Caron-Flinterman, J. E. Broerse, and J. F. Bunders. The experiential knowledge of patients: A new resource for biomedical research? *Social Science and Medicine*, 60(11):2575–2584, 2005. ISSN 02779536. doi: 10.1016/j.socscimed.2004.11.023.

[50] P. Carter, R. Beech, D. Coxon, M. J. Thomas, and C. Jinks. Mobilising the experiential knowledge of clinicians, patients and carers for applied health-care research. *Contemporary Social Science*, 8(3):307–320, 2013. ISSN 21582041. doi: 10.1080/21582041.2013.767468.

[51] P. G. Casali, A. Le Cesne, A. P. Velasco, D. Kotasek, P. Rutkowski, P. Hohenberger, E. Fumagalli, I. R. Judson, A. Italiano, H. Gelderblom, A. Adenis, J. T. Hartmann, F. Duffaud, D. Goldstein, J. M. Broto, A. Gronchi, A. P. Dei Tos, S. Marréaud, W. T. Van Der Graaf, J. R. Zalcberg, S. Litière, and J. Y. Blay. Time to Definitive Failure to the First Tyrosine Kinase Inhibitor in Localized GI Stromal Tumors Treated With Imatinib As an Adjuvant: A European Organisation for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group Intergroup Randomized Trial in Collaboration With the Australasian Gastro-Intestinal Trials Group, UNICANCER, French Sarcoma Group, Italian Sarcoma Group, and Spanish Group for Research on Sarcomas. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 33(36):4276–4283, 2015. ISSN 1527-7755. doi: 10.1200/JCO.2015.62.4304.

[52] P. G. Casali, E. Fumagalli, A. Gronchi, J. Zalcberg, D. Kotasek, A. L. Cesne, J. Y. Blay, A. Italiano, P. Reichardt, L. H. Lindner, V. Grünwald, I. R. Judson, W. Van Der Graaf, P. Schöffski, S. Litière, S. Marreaud, S. Leyvraz, A. L. Pousa, S. Sleijfer, J. Verweij, J. M. Kerst, P. Hogendoorn, W. Van Der Graaf, and P. Rutkowski. Ten-year progression-free and overall survival in patients with unresectable or metastatic GI stromal tumors: Long-term analysis of the european organisation for research and treatment of cancer, Italian sarcoma group, and Australasian gastrointestinal tr. *Journal of Clinical Oncology*, 35(15):1713–1720, 2017. ISSN 15277755. doi: 10.1200/JCO.2016.71.0228.

[53] P. G. Casali, N. Abecassis, S. Bauer, R. Biagini, S. Bielack, S. Bonvalot, I. Boukovinas, J. V. Bovee, T. Brodowicz, J. M. Broto, A. Buonadonna, E. De Álava, A. P. Dei Tos, X. G. Del Muro, P. Dileo, M. Eriksson, A. Fedenko, V. Ferraresi, A. Ferrari, S. Ferrari, A. M. Frezza, S. Gasperoni, H. Gelderblom, T. Gil, G. Grignani, A. Gronchi, R. L. Haas, A. Hannu, B. Hassan, P. Hohenberger, R. Issels, H. Joensuu, R. L. Jones, I. Judson, P. Jutte, S. Kaal, B. Kasper, K. Kopeckova, D. A. Krákorová, A. Le Cesne, I. Lugowska, O. Merimsky, M. Montemurro, M. A. Pantaleo, R. Piana, P. Picci, S. Piperno-Neumann, A. L. Pousa, P. Reichardt, M. H. Robinson, P. Rutkowski, A. A. Safwat, P. Schöffski, S. Sleijfer, S. Stacchiotti, K. Sundby Hall, M. Unk, F. Van Coevorden, W. Van Der Graaf, J. Whelan, E. Wardelmann, O. Zaikova, and J. Y. Blay. Gastrointestinal stromal tumours: ESMO-EURACAN Clinical Practice Guidelines

for diagnosis, treatment and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology*, 29(Suppl 4):iv68–iv78, 2018. ISSN 1569-8041. doi: 10.1093/ANNONC/MDY095.

[54] M. Casparie, A. T. Tiebosch, G. Burger, H. Blauwgeers, A. Van De Pol, J. H. Van Krieken, and G. A. Meijer. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cellular oncology : the official journal of the International Society for Cellular Oncology*, 29(1):19–24, 2007. ISSN 1570-5870. doi: 10.1155/2007/971816.

[55] O. Caster, J. Dietrich, M. L. Kürzinger, M. Lerch, S. Maskell, G. N. Norén, S. Tcherny-Lessenot, B. Vroman, A. Wisniewski, and J. van Stekelenborg. Assessment of the Utility of Social Media for Broad-Ranging Statistical Signal Detection in Pharmacovigilance: Results from the WEB-RADR Project. *Drug Safety*, 41(12):1355–1369, 2018. ISSN 11791942. doi: 10.1007/s40264-018-0699-2.

[56] E. M. Castro, T. Van Regenmortel, W. Sermeus, and K. Vanhaecht. Patients' experiential knowledge and expertise in health care: A hybrid concept analysis. *Social Theory and Health*, 17(3):307–330, 2019. ISSN 1477822X. doi: 10.1057/s41285-018-0081-6.

[57] D. Cer, Y. Yang, S.-y. Kong, N. Hua, N. Limtiaco, R. St. John, N. Constant, M. Guajardo-Cespedes, S. Yuan, C. Tar, B. Strope, and R. Kurzweil. Universal sentence encoder for English. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 169–174, Brussels, Belgium, 2018. Association for Computational Linguistics. doi: 10.18653/v1/D18-2029.

[58] N. Cesare, C. Grant, and E. O. Nsoesie. Understanding demographic bias and representation in social media health data. In *WebSci 2019 - Companion of the 11th ACM Conference on Web Science*, pages 7–9, 2019. ISBN 9781450361743. doi: 10.1145/3328413.3328415.

[59] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, and I. Androutsopoulos. Large-Scale Multi-Label Text Classification on EU Legislation. *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, pages 6314–6322, 2019.

[60] W. W. Chapman, W. Bridewell, P. Hanbury, G. F. Cooper, and B. G. Buchanan. A simple algorithm for identifying negated findings and diseases in discharge summaries. *Journal of Biomedical Informatics*, 34(5):301–310, 2001. ISSN 15320464. doi: 10.1006/jbin.2001.1029.

[61] M. E. Charlson, P. Pompei, K. L. Ales, and C. R. MacKenzie. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases*, 40(5):373–383, 1987. ISSN 0021-9681. doi: 10.1016/0021-9681(87)90171-8. URL `https://pubmed.ncbi.nlm.nih.gov/3558716/`.

[62] H. Chen, W. Chen, C. Liu, L. Zhang, J. Su, and X. Zhou. Relational Network for Knowledge Discovery through Heterogeneous Biomedical and Clinical Features. *Nature Publishing Group*, 2016. doi: 10.1038/srep29915.

[63] G. G. Chowdhury. Natural language processing. *Annual Review of Information Science and Technology*, 37(1):51–89, 2005. ISSN 00664200. doi: 10.1002/aris.1440370103.

[64] K. W. Church and W. A. Gale. Probability scoring for spelling correction. *Statistics and Computing*, 1(2):93–103, 1991. ISSN 1573-1375. doi: 10.1007/BF01889984.

[65] E. Clark and K. Araki. Text Normalization in Social Media: Progress, Problems and Applications for a Pre-processing System of Casual English. *Procedia Soc Behav Sci*, 27:2–11, 2011. doi: 10.1016/j.sbspro.2011.10.577. URL `www.urbandictionary.com`.

[66] K. Clark, M.-T. Luong, G. Brain, Q. V. Le Google Brain, and C. D. Manning. ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators. *ArXiv*, 2020. URL `https://arxiv.org/abs/2003.10555v1`.

[67] A. Cocos, A. G. Fiks, and A. J. Masino. Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in Twitter posts. *Journal of the American Medical Informatics Association*, 24(4):813–821, 2017. doi: 10.1093/jamia/ocw180.

[68] D. A. Coll, C. A. Rosen, K. Auborn, W. P. Potsic, and H. L. Bradlow. Treatment of Recurrent Respiratory Papillomatosis With Indole-3Xarbinol. *American Journal of Otolaryngology*, 18(4):283–285, 1997.

[69] J. A. Custers, R. Tielen, J. B. Prins, J. H. De Wilt, M. F. Gielissen, and W. T. Van Der Graaf. Fear of progression in patients with gastrointestinal stromal tumors (GIST): Is extended lifetime related to the Sword of Damocles? *Acta oncologica (Stockholm, Sweden)*, 54(8):1202–1208, 2015. ISSN 1651-226X. doi: 10.3109/0284186X.2014.1003960.

[70] J. D' Souza and V. Ng. Sieve-Based Entity Linking for the Biomedical Domain. In *53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, pages 297–302, 2015.

[71] R. Dagher, M. Cohen, G. Williams, M. Rothmann, J. Gobburu, G. Robbie, A. Rahman, G. Chen, A. Staten, D. Griebel, R. Pazdur, A. D. V. d. Abbeele, E. v. Sonnenberg, and G. D. Demetri. Approval summary: imatinib mesylate in the treatment of metastatic and/or unresectable malignant gastrointestinal stromal tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 8(10):3034–8, 10 2002. ISSN 1078-0432. doi: 10.1158/1078-0432.ccr-06-0858.

[72] X. Dai. Recognizing complex entity mentions: A review and future directions. In *Proceedings of ACL 2018, Student Research Workshop*, pages 37–44. Association for Computational Linguistics, 2018. doi: 10.18653/v1/P18-3006.

[73] G. D'Amato, D. M. Steinert, J. C. McAuliffe, and J. C. Trent. Update on the biology and therapy of gastrointestinal stromal tumors. *Cancer control : journal of the Moffitt Cancer Center*, 12(1):44–56, 2005. ISSN 1073-2748. doi: 10.1177/107327480501200106.

[74] H. T. Dang. Overview of DUC 2006. In *Proceedings of HLT-NAACL 2006*, pages 1–12, 2006.

[75] K. P. Davison, J. W. Pennebaker, and S. S. Dickerson. Who talks? The social psychology of illness support groups. *American Psychologist*, 55(2):205–217, 2000. ISSN 0003066X. doi: 10.1037/0003-066X.55.2.205.

[76] M. de Boer, A. Dirkson, G. van Oortmerssen, S. Verberne, and W. Kraaij. The Patient Forum Miner: Text mining for patient communities. In *Proceedings of the 17th Dutch-Belgian Information Retrieval Workshop*, pages 25–26, 2018.

[77] J. De Langen, F. Van Hunsel, A. Passier, L. De Jong-Van Den Berg, and K. Van Grootheest. Adverse drug reaction reporting by patients in the Netherlands: Three years of experience. *Drug Safety*, 31(6):515–524, 2008. ISSN 01145916. doi: 10.2165/00002018-200831060-00006.

[78] G. D. Demetri, M. von Mehren, C. D. Blanke, A. D. Van den Abbeele, B. Eisenberg, P. J. Roberts, M. C. Heinrich, D. A. Tuveson, S. Singer, M. Janicek, J. A. Fletcher, S. G. Silverman, S. L. Silberman, R. Capdeville, B. Kiese, B. Peng, S. Dimitrijevic, B. J. Druker, C. Corless, C. D. Fletcher, and H. Joensuu. Efficacy and Safety of Imatinib Mesylate in Advanced Gastrointestinal Stromal Tumors. *New England Journal of Medicine*, 347(7):472–480, 2002. ISSN 0028-4793. doi: 10.1056/NEJMoa020461.

[79] G. D. Demetri, A. T. van Oosterom, C. R. Garrett, M. E. Blackstein, M. H. Shah, J. Verweij, G. McArthur, I. R. Judson, M. C. Heinrich, J. A. Morgan, J. Desai, C. D. Fletcher, S. George, C. L. Bello, X. Huang, C. M. Baum, and P. G. Casali. Efficacy and safety of sunitinib in patients with advanced gastrointestinal stromal tumour after failure of imatinib: a randomised controlled trial. *The Lancet*, 368(9544):1329–1338, 2006. ISSN 0140-6736. doi: 10.1016/S0140-6736(06)69446-4.

[80] G. D. Demetri, P. Reichardt, Y.-K. Kang, J.-Y. Blay, H. Joensuu, R. G. Maki, P. Rutkowski, P. Hohenberger, H. Gelderblom, M. G. Leahy, M. von Mehren, P. Schoffski, M. E. Blackstein, A. Le Cesne, G. Badalamenti, J.-M. Xu, T. Nishida, D. Laurent, I. Kuss, and P. G. Casali. Randomized phase III trial of regorafenib in patients (pts) with metastatic and/or unresectable gastrointestinal stromal tumor (GIST) progressing despite prior treatment with at least imatinib (IM) and sunitinib (SU): GRID trial. *Journal of Clinical Oncology*, 30(18_suppl):LBA10008–LBA10008, 2012. ISSN 0732-183X. doi: 10.1200/jco.2012.30.18_suppl.lba10008.

[81] G. D. Demetri, P. Reichardt, Y.-K. Kang, J.-Y. Blay, P. Rutkowski, H. Gelderblom, P. Hohenberger, M. Leahy, M. von Mehren, H. Joensuu, G. Badalamenti, M. Blackstein, A. Le Cesne, P. Schöffski, R. G. Maki, S. Bauer, B. B. Nguyen, J. Xu, T. Nishida, J. Chung, C. Kappeler, I. Kuss, D. Laurent, P. G. Casali, and GRID

study investigators. Efficacy and safety of regorafenib for advanced gastrointestinal stromal tumours after failure of imatinib and sunitinib (GRID): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *The Lancet*, 381(9863): 295–302, 2013. ISSN 01406736. doi: 10.1016/S0140-6736(12)61857-1.

[82] D. den Hollander, A. R. Dirkson, S. Verberne, W. Kraaij, G. van Oortmerssen, H. Gelderblom, A. Oosten, A. K. L. Reyners, N. Steeghs, W. T. A. van der Graaf, I. M. E. Desar, and O. Husson. Symptoms reported by gastrointestinal stromal tumour (GIST) patients on imatinib treatment: combining questionnaire and forum data. *Supportive Care in Cancer*, mar 2022. ISSN 0941-4355. doi: 10.1007/S00520-022-06929-3. URL https://link.springer.com/10.1007/s00520-022-06929-3.

[83] L. Derczynski, E. Nichols, and M. van Erp. Results of the WNUT2017 Shared Task on Novel and Emerging Entity Recognition. In *Proceedings ofthe 3rd Workshop on Noisy User-generated Text*, pages 140–147. Association for Computational Linguistics, 2017.

[84] 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)*, pages 4171–4186, Minneapolis, Minnesota, 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1423.

[85] M. Di Maio, E. Basch, J. Bryce, and F. Perrone. Patient-reported outcomes in the evaluation of toxicity of anticancer treatments. *Nature reviews. Clinical oncology*, 13(5):319–325, may 2016. ISSN 1759-4782. doi: 10.1038/NRCLINONC.2015.222.

[86] G.-A. Dima, D.-C. Cercel, and M. Dascalu. Transformer-based Multi-Task Learning for Adverse Effect Mention Analysis in Tweets. In *Proceedings of the Sixth Social Media Mining for Health Workshop 2021*, pages 44–51, 2021. doi: 10.18653/v1/2021.smm4h-1.7.

[87] A. Dirkson, S. Verberne, and W. Kraaij. Narrative Detection in Online Patient Communities. In A. Jorge, R. Campos, A. Jatowt, and S. Bhatia, editors, *Proceedings of the Text2StoryIR'19 Workshop*. CEUR-WS, 2019. URL http://ceur-ws.org/Vol-2342/paper3.pdf.

[88] A. Dirkson, S. Verberne, W. Kraaij, G. V. Oortmerssen, and H. Gelderblom. Automated gathering of real-world evidence from online patient fora can complement pharma. *Scientific Reports*, 2022. doi: 10.1038/s41598-022-13894-8.

[89] Y. Doval Mosquera, J. Vilares, and C. Gómez-Rodríguez. Lysgroup: Adapting a spanish microtext normalization system to english. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 99–105, Beijing, China, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4315.

[90] R. I. Doğan, R. Leaman, and Z. Lu. NCBI disease corpus: A resource for disease name recognition and concept normalization. *Journal of Biomedical Informatics*, 47:1–10, 2014. ISSN 15320464. doi: 10.1016/j.jbi.2013.12.006.

[91] M. Dredze, D. A. Broniatowski, and K. M. Hilyard. Zika vaccine misconceptions: A social media analysis. *Vaccine*, 34(30):3441–2, 2016. ISSN 1873-2518. doi: 10.1016/j.vaccine.2016.05.008.

[92] ECSA. Ten principles of Citizen Science. Technical report, European Citizen Science Association, 2015. URL http://doi.org/10.17605/OSF.IO/XPR2N.

[93] G. I. Ector, R. P. Hermens, and N. M. Blijlevens. Filling the gaps of patient information and comprehension. *Current opinion in oncology*, 32(4):262–268, 2020. ISSN 1531-703X. doi: 10.1097/CCO.0000000000000633.

[94] I. Eekhout, M. A. Van De Wiel, and M. W. Heymans. Methods for significance testing of categorical covariates in logistic regression models after multiple imputation: Power and applicability analysis. *BMC Medical Research Methodology*, 17(1):1–12, 2017. ISSN 14712288. doi: 10.1186/S12874-017-0404-7/TABLES/4. URL https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0404-7.

[95] I. Eerola. Rare Diseases remains a major unmet medical need. Technical report, European Commission, Brussels, 2017. URL https://doi.org/10.2777/76751.

[96] F. Efficace, G. Rosti, N. Aaronson, F. Cottone, E. Angelucci, S. Molica, M. Vignetti, F. Mandelli, and M. Baccarani. Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia. *Haematologica*, 99(4):788–793, 2014. ISSN 1592-8721. doi: 10.3324/HAEMATOL.2013.093724.

[97] J. Eisenstein. What to do about bad language on the internet. In *Proceedings of NAACL-HLT 2013*, pages 9–14. Association for Computational Linguistics, 2013.

[98] I. A. Eland, K. J. Belton, A. C. Van Grootheest, A. P. Meiners, M. D. Rawlins, and B. H. Stricker. Attitudinal survey of voluntary reporting of adverse drug reactions. *British Journal of Clinical Pharmacology*, 48(4):623–627, 1999. ISSN 03065251. doi: 10.1046/j.1365-2125.1999.00060.x.

[99] J. L. Elman. Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1):71–99, 1993.

[100] EORTC Quality of Life Group. EORTC QLQ-30 Scoring Manual, 2001. URL https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf.

[101] European Medicine Agency. Summary of Product Characteristics Imatinib. Technical report, European Medicine Agency. URL https://www.ema.europa.eu/en/documents/product-information/glivec-epar-product-information_en.pdf.

[102] European Medicine Agency. Guideline on good pharmacovigilance practices (GVP) - Annex I - Definitions (Rev 4). Technical report, European Medicine Agency, 2017. URL https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-annex-i-definitions-rev-4_en.pdf.

[103] European Medicine Agency. Ayvakyt, 2020. URL https://www.ema.europa.eu/en/medicines/human/EPAR/ayvakyt.

[104] N. Fareed, C. M. Swoboda, P. Jonnalagadda, and T. R. Huerta. Persistent digital divide in health-related internet use among cancer survivors: findings from the Health Information National Trends Survey, 2003-2018. *Journal of cancer survivorship: research and practice*, 15(1):87–98, 2021. ISSN 1932-2267. doi: 10.1007/S11764-020-00913-8.

[105] L. Fauske, I. Hompland, G. Lorem, H. Bondevik, and Ø. S. Bruland. Perspectives on treatment side effects in patients with metastatic gastrointestinal stromal tumour: a qualitative study. *Clinical Sarcoma Research 2019 9:1*, 9(1):1–8, 2019. ISSN 2045-3329. doi: 10.1186/S13569-019-0116-3.

[106] P. Fayers and A. Bottomley. Quality of life research within the EORTC - The EORTC QLQ-C30. *European Journal of Cancer*, 38:S125–133, 2002. ISSN 09598049. doi: 10.1016/s0959-8049(01)00448-8.

[107] R. Feldman, J. Sanger, and C. U. Press. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press, 2007. ISBN 9780521836579.

[108] T. Ferguson. The first generation of e-patients. *BMJ*, 328(7449):1148–1149, 2004. ISSN 0959-8138. doi: 10.1136/bmj.328.7449.1148.

[109] T. Ferguson and B. Kelly. E-patients prefer egroups to doctors for 10 of 12 aspects of health care., 1991. URL http://www.fergusonreport.com/articles/fr039905.htm.

[110] P. Fivez, S. Šuster, and W. Daelemans. Unsupervised context-sensitive spelling correction of clinical free-text with word and character n-gram embeddings. In *BioNLP 2017*, pages 143–148, Vancouver, Canada,, 2017. Association for Computational Linguistics. doi: 10.18653/v1/W17-2317.

[111] J. Fjeld, N. Achten, H. Hilligoss, A. C. Nagy, and M. Srikumar. Principled Artificial intelligence: Mapping consensus in Ethical and Rights-based Approaches to Principles for AI. Technical report, Berkman Klein Center for Internet & Society, 2020.

[112] R. Ginn, P. Pimpalkhute, A. Nikfarjam, A. Patki, K. O 'connor, A. Sarker, K. Smith, and G. Gonzalez. Mining Twitter for Adverse Drug Reaction Mentions: A Corpus and Classification Benchmark. In *Proceedings of the fourth workshop on building and evaluating resources for health and biomedical text processing*, 2014.

[113] W. G. Goettsch, S. D. Bos, N. Breekveldt-Postma, M. Casparie, R. M. Herings, and P. C. Hogendoorn. Incidence of gastrointestinal stromal tumours is underestimated: Results of a nation-wide study. *European Journal of Cancer*, 41(18):2868–2872, 2005. ISSN 0959-8049. doi: 10.1016/J.EJCA.2005.09.009.

[114] S. Golder, G. Norman, and Y. K. Loke. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *British Journal of Clinical Pharmacology*, 80(4):878–888, 2015. ISSN 13652125. doi: 10.1111/bcp.12746.

[115] S. Golder, K. Smith, K. O'Connor, R. Gross, S. Hennessy, and G. Gonzalez-Hernandez. A Comparative View of Reported Adverse Effects of Statins in Social Media, Regulatory Data, Drug Information Databases and Systematic Reviews. *Drug Safety*, pages 1–13, 2020. ISSN 11791942. doi: 10.1007/s40264-020-00998-1.

[116] G. Gonzalez-Hernandez, A. Sarker, K. O 'Connor, and G. Savova. Capturing the Patient's Perspective : a Review of Advances in Natural Language Processing of Health-Related Text. *Yearbook of medical informatics*, pages 214–217, 2017. doi: 10.15265/IY-2017-029.

[117] I. J. Goodfellow, J. Shlens, and C. Szegedy. Explaining and Harnessing Adversarial Examples. In *ICLR 2015*, 2015.

[118] R. C. Griggs, M. Batshaw, M. Dunkle, R. Gopal-Srivastava, E. Kaye, J. Krischer, T. Nguyen, K. Paulus, and P. A. Merkel. Clinical research for rare disease: Opportunities, challenges, and solutions. *Molecular Genetics and Metabolism*, 96 (1):20–26, 2009. ISSN 10967192. doi: 10.1016/j.ymgme.2008.10.003.

[119] Y. Gu, R. Tinn, H. Cheng, M. Lucas, N. Usuyama, X. Liu, T. Naumann, J. Gao, and H. Poon. Domain-specific language model pretraining for biomedical natural language processing. *ACM Transactions on Computing for Healthcare (HEALTH)*, 3 (1):1–23, 2021. doi: 10.1145/3458754.

[120] B. Han and T. Baldwin. Lexical Normalization for Social Media Text. *ACM Trans. Intell. Syst. Technol. Article*, 4(5), 2013. doi: 10.1145/2414425.2414430.

[121] B. Han, P. Cook, and T. Baldwin. Automatically constructing a normalisation dictionary for microblogs. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 421–432. Association for Computational Linguistics, 2012.

[122] J. Y. Han, J.-H. Kim, H. J. Yoon, M. Shim, F. M. McTavish, and D. H. Gustafson. Social and Psychological Determinants of Levels of Engagement With an Online Breast Cancer Support Group: Posters, Lurkers, and Nonusers. *Journal of Health Communication*, 17(3):356–371, 2012. doi: 10.1080/10810730.2011.585696.

[123] J. Y. Han, J. Hou, E. Kim, and D. H. Gustafson. Lurking as an Active Participation Process: A Longitudinal Investigation of Engagement with an Online Cancer Support Group. *Health Communication*, 29(9):911–923, 2014. doi: 10.1080/10410236.2013.816911.

[124] S. Han, T. Tran, A. Rios, and R. Kavuluru. Team UKNLP: Detecting ADRs, Classifying Medication Intake Messages, and Normalizing ADR Mentions on Twitter. In *Proceedings of the Second Workshop on Social Media Mining for Health Applications (SMM4H)*, 2017.

[125] E. Hargittai. Whose Space? Differences Among Users and Non-Users of Social Network Sites. *Journal of Computer-Mediated Communication*, 13(1):276–297, 2007. ISSN 10836101. doi: 10.1111/j.1083-6101.2007.00396.x.

[126] E. Hargittai. Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites. *The ANNALS of the American Academy of Political and Social Science*, 659(1):63–76, 2015. ISSN 0002-7162. doi: 10.1177/0002716215570866.

[127] E. Hargittai. Potential Biases in Big Data: Omitted Voices on Social Media. *Social Science Computer Review*, 38(1):10–24, 2020. doi: 10.1177/0894439318788322.

[128] R. Harpaz, W. DuMouchel, N. H. Shah, D. Madigan, P. Ryan, and C. Friedman. Novel Data-Mining Methodologies for Adverse Drug Event Discovery and Analysis. *Clinical Pharmacology & Therapeutics*, 91(6):1010–1021, 2012. ISSN 0009-9236. doi: 10.1038/clpt.2012.50.

[129] A. Hartzler and W. Pratt. Managing the personal side of health: how patient expertise differs from the expertise of clinicians. *Journal of medical Internet research*, 13(3): e62, 2011. ISSN 1438-8871. doi: 10.2196/jmir.1728.

[130] L. Hazell and S. A. W. Shakir. Under-Reporting of Adverse Drug Reactions A Systematic Review. *Drug Safety*, 29(5):385–396, 2006. doi: 10.2165/00002018-200629050-00003.

[131] H. E. Heemstra, S. van Weely, H. A. Büller, H. G. Leufkens, and R. L. de Vrueh. Translation of rare disease research into orphan drug development: disease matters. *Drug Discovery Today*, 14(23-24):1166–1173, 2009. ISSN 13596446. doi: 10.1016/j.drudis.2009.09.008.

[132] G. Heigold, G. Neumann, and J. van Genabith. How Robust Are Character-Based Word Embeddings in Tagging and MT Against Wrod Scramlbing or Randdm Nouse? In *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Papers)*, pages 68–80, 2018.

[133] High-Level Expert Group on AI (AI HLEG). Ethics guidelines for trustworthy AI. Technical report, European Commission, 2019. URL `https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html`.

[134] F. Hill, R. Reichart, and A. Korhonen. SimLex-999: Evaluating Semantic Models With (Genuine) Similarity Estimation. *Computational Linguistics*, 41(4):665–695, 2015. doi: 10.1162/COLI.

[135] J. Hong, T. Kim, H. Lim, and J. Choo. AVocaDo: Strategy for Adapting Vocabulary to Downstream Domain. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 4692–4700, 2021.

[136] M. Honnibal and I. Montani. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear, 2017.

[137] J. Howard and S. Ruder. Universal language model fine-tuning for text classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 328–339, Melbourne, Australia, 2018. Association for Computational Linguistics.

[138] G. Hripcsak and D. F. Heitjan. Measuring agreement in medical informatics reliability studies. *Journal of Biomedical Informatics*, 35(2):99–110, 2002. ISSN 1532-0464. doi: 10.1016/S1532-0464(02)00500-2.

[139] Y.-L. Hsieh, M. Cheng, D.-C. Juan, W. Wei, W.-L. Hsu, and C.-J. Hsieh. On the Robustness of Self-Attentive Models. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1520–1529, 2019. doi: 10.18653/v1/P19-1147.

[140] Y. Hu, X. Jing, Y. Ko, and J. T. Rayz. Misspelling Correction with Pre-trained Contextual Language Model. *arXiv*, 2021. URL `https://arxiv.org/abs/2101.03204`.

[141] X. Huang, M. C. Smith, M. Paul, D. Ryzhkov, S. Quinn, D. Broniatowski, and M. Dredze. Examining Patterns of Influenza Vaccination in Social Media. In *AAAI Joint Workshop on Health Intelligence (W3PHIAI)*, 2017.

[142] Hyeoneui, J. Mentzer, and R. Taira. Developing a Physical Activity Ontology to Support the Interoperability of Physical Activity Data. *J Med Internet Res 2019;21(4):e12776 https://www.jmir.org/2019/4/e12776*, 21(4):e12776, 2019. doi: 10.2196/12776.

[143] D. Jin, Z. Jin, J. T. Zhou, and P. Szolovits. Is bert really robust? a strong baseline for natural language attack on text classification and entailment. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):8018–8025, 2020. doi: 10.1609/aaai.v34i05.6311.

[144] N. Jin. NCSU-SAS-ning: Candidate generation and feature engineering for supervised lexical normalization. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 87–92, Beijing, China, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4313.

[145] H. Joensuu, J. C. Trent, and P. Reichardt. Practical management of tyrosine kinase inhibitor-associated side effects in GIST. *Cancer treatment reviews*, 37(1):75–88, 2011. ISSN 1532-1967. doi: 10.1016/J.CTRV.2010.04.008.

[146] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark. Mimic-iii, a freely accessible critical care database. *Scientific data*, 3(1):1–9, 2016.

[147] C. Johnson, N. Aaronson, J. M. Blazeby, A. Bottomley, P. Fayers, M. Koller, D. Kuliś, J. Ramage, M. Sprangers, G. Velikova, and T. Young. EORTC QUALITY OF LIFE GROUP Guidelines for Developing Questionnaire Modules. Technical report, European Organisation for Research and Treatment of Cancer, 2011. URL https://www.eortc.org/app/uploads/sites/2/2018/02/guidelines_for_developing_questionnaire-_final.pdf.

[148] M. O. Johnson and T. B. Neilands. Coping with HIV treatment side effects: Conceptualization, measurement, and linkages. *AIDS and Behavior*, 11(4):575–585, 2007. doi: 10.1007/S10461-007-9229-4.

[149] D. Jurafsky and J. H. Martin. *Speech and Language Processing*. Pearson Education, 3rd edition, 2021.

[150] K. R. Kanakarajan, B. Kundumani, and M. Sankarasubbu. BioELECTRA:Pretrained Biomedical text Encoder using Discriminators. In *Proceedings of the BioNLP 2021 workshop*, pages 143–154, 2021.

[151] S. Karimi, A. Metke-Jimenez, M. Kemp, and C. Wang. Cadec: A corpus of adverse drug event annotations. *Journal of Biomedical Informatics*, 55:73–81, 2015. ISSN 1532-0464. doi: 10.1016/J.JBI.2015.03.010.

[152] P. Karisani and E. Agichtein. Did you really just have a heart attack?: Towards robust detection of personal health mentions in social media. In *Proceedings of the 2018 World Wide Web Conference*, pages 137–146, 2018. ISBN 978-1-4503-5639-8. doi: 10.1145/3178876.3186055.

[153] R. Kaur, J. A. Ginige, O. Obst, and A. Ginige. A Systematic Literature Review of Automated ICD Coding and Classification Systems using Discharge Summaries. *ArXiv (preprint)*, 2021. URL https://arxiv.org/abs/2107.10652.

[154] S. Khosla, R. White, J. Medina, M. Ouwens, C. Emmas, T. Koder, G. Male, and S. Leonard. Real world evidence (RWE) - a disruptive innovation or the quiet evolution of medical evidence generation? *F1000Research*, 7:111, 2018. ISSN 20461402. doi: 10.12688/f1000research.13585.2.

[155] I. Kickbusch and D. Gleicher. Governance for Health in the 21st Century. Technical report, World Health Organization (WHO), Copenhagen, 2012. URL https://www.euro.who.int/en/publications/abstracts/governance-for-health-in-the-21st-century.

[156] T. Kikuchi, T. Suzuki, H. Uchida, K. Watanabe, and M. Mimura. Coping strategies for antidepressant side effects: An Internet survey. *Journal of Affective Disorders*, 143 (1-3):89–94, dec 2012. ISSN 0165-0327. doi: 10.1016/J.JAD.2012.04.039.

[157] N. Kingod, B. Cleal, A. Wahlberg, and G. R. Husted. Online peer-to-peer communities in the daily lives of people with chronic illness: A qualitative systematic review. *Qualitative Health Research*, 27(1):89–99, 2017. doi: 10.1177/1049732316680203.

[158] A. Kinsora, K. Barron, Q. Mei, and V. G. V. Vydiswaran. Creating a Labeled Dataset for Medical Misinformation in Health Forums. In *IEEE International Conference on Healthcare Informatics*, 2017. doi: 10.1109/ICHI.2017.93.

[159] A. Klein, A. Magge, K. O'Connor, H. Cai, D. Weissenbacher, and G. Gonzalez-Hernandez. A chronological and geographical analysis of personal reports of covid-19 on twitter. *medRxiv (preprint)*, 2020. doi: 10.1101/2020.04.19.20069948.

[160] D. C. Klonoff, A. Gutierrez, A. Fleming, and D. Kerr. Real-World Evidence Should Be Used in Regulatory Decisions About New Pharmaceutical and Medical Device Products for Diabetes. *Journal of Diabetes Science and Technology*, 13(6):995–1000, 2019. ISSN 19322968. doi: 10.1177/1932296819839996.

[161] P. G. Kluetz, A. Slagle, E. J. Papadopoulos, L. L. Johnson, M. Donoghue, V. E. Kwitkowski, W. H. Chen, R. Sridhara, A. T. Farrell, P. Keegan, G. Kim, and R. Pazdur. Focusing on Core Patient-Reported Outcomes in Cancer Clinical Trials: Symptomatic Adverse Events, Physical Function, and Disease-Related Symptoms. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 22(7):1553–1558, 2016. ISSN 1557-3265. doi: 10.1158/1078-0432.CCR-15-2035.

[162] I. Koiranen, ·. Teo Keipi, ·. A. Koivula, and P. Räsänen. Changing patterns of social media use? A population-level study of Finland. *Universal Access in the Information Society*, 19:603–617, 2020. doi: 10.1007/s10209-019-00654-1.

[163] R. Kraut, S. Kiesler, B. Boneva, J. Cummings, V. Helgeson, and A. Crawford. Internet Paradox Revisited. *Journal of Social Issues*, 58(1):49–74, 2002.

[164] K. Kukich. Techniques for automatically correcting words in text. *ACM Comput. Surv.*, 24:377–439, 1992.

[165] D. Kulis, A. Bottomley, C. Whittaker, L. van de Poll-Franse, A. Darlington, B. Holzner, M. Koller, J. Reijneveld, K. Tomaszewski, and M. Grønvold. The Use of The Eortc Item Library To Supplement Eortc Quality of Life Instruments. *Value in Health*, 20 (9):A775, 2017. ISSN 10983015. doi: 10.1016/J.JVAL.2017.08.2236.

[166] A. Kumar, P. Makhija, and A. Gupta. Noisy Text Data: Achilles' Heel of BERT. In *Proceedings of the 2020 EMNLP Workshop W-NUT: The Sixth Workshop on Noisy User-generated Text*, pages 16–21. Association for Computational Linguistics, 2020.

[167] K. Kvarnström, A. Westerholm, M. Airaksinen, and H. Liira. Factors Contributing to Medication Adherence in Patients with a Chronic Condition: A Scoping Review of Qualitative Research. *Pharmaceutics 2021, Vol. 13, Page 1100*, 13(7):1100, jul 2021. doi: 10.3390/PHARMACEUTICS13071100.

[168] K. H. Lai, M. Topaz, F. R. Goss, and L. Zhou. Automated misspelling detection and correction in clinical free-text records. *Journal of Biomedical Informatics*, 55:188–195, 2015. doi: 10.1016/j.jbi.2015.04.008.

[169] A. Lamb, M. J. Paul, and M. Dredze. Separating Fact from Fear: Tracking Flu Infections on Twitter. In *Proceedings of NAACL-HLT*, pages 789–795, 2013.

[170] J. B. Lamy, A. Venot, and C. Duclos. PyMedTermino: An open-source generic API for advanced terminology services. *Studies in Health Technology and Informatics*, 210: 924–928, 2015. ISSN 18798365. doi: 10.3233/978-1-61499-512-8-924.

[171] J. Lardon, R. Abdellaoui, F. Bellet, H. Asfari, J. Souvignet, N. Texier, M. C. Jaulent, M. N. Beyens, A. Burgun, and C. Bousquet. Adverse drug reaction identification and extraction in social media: A scoping review. *Journal of Medical Internet Research*, 17(7):1–16, 2015. ISSN 14388871. doi: 10.2196/jmir.4304.

[172] Q. Le and T. Mikolov. Distributed Representations of Sentences and Documents. In *Proceedings of the 31st international conference on machine learning*, 2014.

[173] R. Leaman, L. Wojtulewicz, R. Sullivan, A. Skariah, J. Yang, and G. Gonzalez. Towards Internet-age pharmacovigilance: Extracting adverse drug reactions from user posts to health-related social networks. *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing, ACL 2010*, pages 117–125, 2010.

[174] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4):1234–1240, 2019. doi: 10.1093/bioinformatics/btz682.

[175] S. Leeman-Munk, J. Lester, and J. Cox. NCSU_SAS_SAM: Deep encoding and reconstruction for normalization of noisy text. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 154–161, Beijing, China, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4323.

[176] M. Lengsavath, A. Dal Pra, A. M. de Ferran, S. Brosch, L. Härmark, V. Newbould, and S. Goncalves. Social Media Monitoring and Adverse Drug Reaction Reporting in Pharmacovigilance: An Overview of the Regulatory Landscape. *Therapeutic Innovation and Regulatory Science*, 51(1):125–131, 2017. ISSN 21684804. doi: 10.1177/2168479016663264.

[177] D. D. Lewis, Y. Yang, T. G. Rose, and F. Li. RCV1: A New Benchmark Collection for Text Categorization Research. *Journal of Machine Learning Research*, 5:361–397, 2004.

[178] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer, and F. Ai. 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*, pages 7871–7880. Association for Computational Linguistics (ACL), jul 2020. doi: 10.18653/V1/2020.ACL-MAIN.703.

[179] J. Li, Y. Sun, R. J. Johnson, D. Sciaky, C.-H. Wei, R. Leaman, A. P. Davis, C. J. Mattingly, T. C. Wiegers, and Z. Lu. BioCreative V CDR task corpus: a resource for chemical disease relation extraction. *Database*, 2016:68, 2016. doi: 10.1093/database/baw068.

[180] L. Li, R. Ma, Q. Guo, X. Xue, and X. Qiu. Bert-attack: Adversarial attack against BERT using BERT. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pages 6193–6202, 2020. doi: 10.18653/v1/2020.emnlp-main.500.

[181] Q. Li, Q. Zhang, and L. Si. eventAI at SemEval-2019 Task 7: Rumor Detection on Social Media by Exploiting Content, User Credibility and Propagation Information. In *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019)*, pages 855–859, 2019. doi: 10.18653/v1/s19-2148.

[182] X. Li, W. Gao, S. Feng, Y. Zhang, and D. Wang. Boundary detection with BERT for span-level emotion cause analysis. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 676–682. Association for Computational Linguistics, 2021. doi: 10.18653/v1/2021.findings-acl.60.

[183] Z. Li, Z. Yang, L. Luo, Y. Xiang, and H. Lin. Exploiting adversarial transfer learning for adverse drug reaction detection from texts. *Journal of Biomedical Informatics*, page 103431, 2020. doi: 10.1016/j.jbi.2020.103431.

[184] N. Limsopatham and N. Collier. Normalising medical concepts in social media texts by learning semantic representation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1014–1023, Berlin, Germany, 2016. Association for Computational Linguistics. doi: 10.18653/v1/P16-1096.

[185] B. Y. Lin, W. Gao, J. Yan, R. Moreno, and X. Ren. RockNER: A Simple Method to Create Adversarial Examples for Evaluating the Robustness of Named Entity Recognition Models. In *Proceedings ofthe 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3728–3737, 2021.

[186] R. J. A. Little and D. B. Rubin. *Statistical Analysis with Missing Data*. John Wiley & Sons, Ltd, New York City, USA, 1986.

[187] F. Liu, Z. Meng, M. Basaldella, and N. Collier. Self-Alignment Pretraining for Biomedical Entity Representations. In *Proceedings of the 2021 Conference ofthe North American Chapter ofthe Association for Computational Linguistics*, pages 4228–4238, 2021.

[188] X. Liu and H. Chen. Azdrugminer: an information extraction system for mining patient-reported adverse drug events in online patient forums. In *International conference on smart health*, pages 134–150. Springer, 2013. doi: 10.1007/978-3-642-39844-5_16.

[189] E. Lopez-Gonzalez, M. T. Herdeiro, and A. Figueiras. Determinants of under-reporting of adverse drug reactions: A systematic review. *Drug Safety*, 32(1):19–31, 2009. ISSN 01145916. doi: 10.2165/00002018-200932010-00002.

[190] M. Lui and T. Baldwin. langid.py: An Off-the-shelf Language Identification Tool. In *Proceedings of the 50th annual meeting of the association of computational linguistics*, pages 25–30, 2012.

[191] X. Ma, P. Xu, Z. Wang, R. Nallapati, and B. Xiang. Domain Adaptation with BERT-based Domain Classification and Data Selection. In *Proceedings of the 2nd Workshop on Deep Learning Approaches for Low-Resource NLP (DeepLo 2019)*, pages 76–83. Association for Computational Linguistics (ACL), 2019. ISBN 9781950737789. doi: 10.18653/V1/D19-6109.

[192] J. B. Macdonald, B. Macdonald, L. E. Golitz, P. LoRusso, and A. Sekulic. Cutaneous adverse effects of targeted therapies: Part I: Inhibitors of the cellular membrane. *Journal of the American Academy of Dermatology*, 72(2):203–218, 2015. ISSN 1097-6787. doi: 10.1016/J.JAAD.2014.07.032.

[193] A. Magge, A. Z. Klein, A. Miranda-Escalada, M. A. Al-Garadi, I. Alimova, Z. Miftahutdinov, S. Lima López, I. Flores, K. O'connor, D. Weissenbacher, E. Tutubalina, J. M. Banda, M. Krallinger, and G. Gonzalez-Hernandez. Overview of the Sixth Social Media Mining for Health Applications (#SMM4H) Shared Tasks at NAACL 2021. In *Proceedings of the Sixth Social Media Mining for Health Workshop 2021*, pages 21–32, 2021. doi: 10.18653/v1/2021.smm4h-1.4.

[194] A. Magge, E. Tutubalina, Z. Miftahutdinov, I. Alimova, A. Dirkson, S. Verberne, D. Weissenbacher, and G. Gonzalez-Hernandez. DeepADEMiner: a deep learning pharmacovigilance pipeline for extraction and normalization of adverse drug event mentions on Twitter. *Journal of the American Medical Informatics Association*, 2021. doi: 10.1093/jamia/ocab114.

[195] A. S. Maiya. ktrain: A low-code library for augmented machine learning. *arXiv*, 2020. URL `https://arxiv.org/pdf/2004.10703.pdf`.

[196] V. Malykh, V. Logacheva, and T. Khakhulin. Robust Word Vectors: Context-Informed Embeddings for Noisy Texts. In *Proceedings of the 2018 EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-generated Text*, pages 54–63, 2018.

[197] J. J. Mao, A. Chung, A. Benton, S. Hill, L. Ungar, C. E. Leonard, S. Hennessy, and J. H. Holmes. Online discussion of drug side effects and discontinuation among breast cancer survivors. *Pharmacoepidemiology and drug safety*, 22(3):256–62, 2013. ISSN 1099-1557. doi: 10.1002/pds.3365.

[198] S. Mariani, E. Abruzzese, S. Basciani, D. Fiore, A. Persichetti, M. Watanabe, G. Spera, and L. Gnessi. Reversible hair depigmentation in a patient treated with imatinib. *Leukemia research*, 35(6), 2011. ISSN 1873-5835. doi: 10.1016/J.LEUKRES.2010.11.028.

[199] K. Y. McKenna and J. A. Bargh. Coming Out in the Age of the Internet: Identity "Demarginalization" Through Virtual Group Participation. *Journal of Personality and Social Psychology*, 75(3):681–694, 1998. doi: 10.1037/0022-3514.75.3.681.

[200] S. Merity, C. Xiong, J. Bradbury, and R. Socher. Pointer Sentinel Mixture Models. In *ICLR 2017*, 2017.

[201] S. Merity, N. S. Keskar, and R. Socher. Regularizing and Optimizing LSTM Language Models. In *ICLR 2018*, 2018.

[202] M. Merolli, K. Gray, and F. Martin-Sanchez. Health outcomes and related effects of using social media in chronic disease management: A literature review and analysis of affordances. *Journal of Biomedical Informatics*, 46(6):957–969, 12 2013. ISSN 1532-0464. doi: 10.1016/J.JBI.2013.04.010.

[203] A. Metke-Jimenez and S. Karimi. Concept Extraction to Identify Adverse Drug Reactions in Medical Forums: A Comparison of Algorithms. *ArXiv*, 2015. URL `http://arxiv.org/abs/1504.06936`.

[204] A. Metke-Jimenez, S. Karimi, and C. Paris. Evaluation of Text-Processing Algorithms for Adverse Drug Event Extraction from Social Media. In *SoMeRA '14 Proceedings of the first international workshop on Social media retrieval and analysis*, pages 15–20, 2014. doi: 10.1145/2632188.2632200.

[205] P. Michael E. Porter. Perspective - What Is Value in Health Care? *The New England Journal of Medicine*, 363(1):1–3, 2010. ISSN 15334406. doi: 10.1056/NEJMp1011024.

[206] Z. Miftahutdinov and E. Tutubalina. Deep Neural Models for Medical Concept Normalization in User-Generated Texts. In *Student Research Workshop ACL 2019*, 2019. doi: 10.18653/v1/P19-2055.

[207] Z. S. Miftahutdinov, E. V. Tutubalina, and A. E. Tropsha. Identifying disease-related expressions in reviews using conditional random fields. In *Proceedings of the International Conference Dialogue 2017*, 2017.

[208] W. Min and B. Mott. NCSU_SAS_WOOKHEE: A deep contextual long-short term memory model for text normalization. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 111–119, Beijing, China, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4317.

[209] I. Mondal. BBAEG: Towards BERT-based biomedical adversarial example generation for text classification. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5378–5384, Online, 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.naacl-main.423.

[210] D. L. Mowery, B. R. South, L. Christensen, J. Leng, L.-M. Peltonen, S. Salanterä, H. Suominen, D. Martinez, S. Velupillai, N. Elhadad, G. Savova, S. Pradhan, and W. W. Chapman. Normalizing acronyms and abbreviations to aid patient understanding of clinical texts: ShARe/CLEF eHealth Challenge 2013, Task 2. *Journal of biomedical semantics*, 7:43, 2016. ISSN 2041-1480. doi: 10.1186/s13326-016-0084-y.

[211] N. Mrkšić, D. Ó Séaghdha, B. Thomson, M. Gašíc, L. Rojas-Barahona, P.-H. Su, D. Vandyke, T.-H. Wen, and S. Young. Counter-fitting Word Vectors to Linguistic Constraints. In *Proceedings of NAACL-HLT 2016*, pages 142–148, 2016.

[212] J. Mullenbach, S. Wiegreffe, J. Duke, J. Sun, and J. Eisenstein. Explainable Prediction of Medical Codes from Clinical Text. In *Proceedings ofNAACL-HLT 2018*, pages 1101–1111, 2018.

[213] B. Muller, B. Sagot, and D. Seddah. Enhancing BERT for Lexical Normalization. In *The 5th Workshop on Noisy User-generated Text (W- NUT)*, 2019.

[214] N. Nakashole. Commonsense about human senses: Labeled data collection processes. In *Proceedings of the First Workshop on Commonsense Inference in Natural Language Processing*, pages 43–52. Association for Computational Linguistics, 2019. doi: 10.18653/v1/D19-6005.

[215] National Cancer Institute. NCI Dictionary of Cancer Terms. URL `https://www.cancer.gov/publications/dictionaries/cancer-terms`.

[216] A. Nikfarjam and G. H. Gonzalez. Pattern mining for extraction of mentions of Adverse Drug Reactions from user comments. In *AMIA Annual Symposium proceedings*, pages 1019–1026. American Medical Informatics Association, 2011.

[217] A. Nikfarjam, A. Sarker, K. O'Connor, R. Ginn, and G. Gonzalez. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *Journal of the American Medical Informatics Association : JAMIA*, 22(3):671–81, 5 2015. ISSN 1527-974X. doi: 10.1093/jamia/ocu041.

[218] J. Niu, Y. Yang, S. Zhang, Z. Sun, and W. Zhang. Multi-task Character-Level Attentional Networks for Medical Concept Normalization. *Neural Processing Letters*, 49(3):1239–1256, 2019. ISSN 1370-4621. doi: 10.1007/s11063-018-9873-x.

[219] P. Norvig. Natural Language Corpus Data. In J. H. Toby Segaran, editor, *Beautiful Data: The Stories Behind Elegant Data Solutions*, pages 219–242. O'Reilly Media, 2009. ISBN 978-0-596-15711-1.

[220] N. F. Noy, N. Griffith, and M. A. Musen. Collecting community-based mappings in an ontology repository. In *Proceedings of the 7th International Conference on The Semantic Web*, page 371–386, 2008. ISBN 9783540885634. doi: 10.1007/978-3-540-88564-1_24.

[221] M. D. T. Nzali, S. Bringay, C. Lavergne, C. Mollevi, and T. Opitz. What patients can tell us: Topic analysis for social media on breast cancer. *JMIR Medical Informatics*, 5(3):1–17, 2017. ISSN 22919694. doi: 10.2196/medinform.7779.

[222] Observational Health Data Sciences and Informatics project. *The book of OHDSI*. Observational Health Data Sciences and Informatics, 2021. ISBN 978-1088855195. URL `https://ohdsi.github.io/TheBookOfOhdsi/`.

[223] D. O'Callaghan, D. Greene, J. Carthy, and P. Cunningham. An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications*, 42 (13):5645–5657, 2015. ISSN 0957-4174. doi: 10.1016/J.ESWA.2015.02.055.

[224] K. O'Connor, P. Pimpalkhute, A. Nikfarjam, R. Ginn, K. L. Smith, and G. Gonzalez. Pharmacovigilance on twitter? Mining tweets for adverse drug reactions. *AMIA Annual Symposium proceedings*, pages 924–933, 2014. ISSN 1942-597X.

[225] B. O'Donovan, R. M. Rodgers, A. R. Cox, and J. Krska. Use of information sources regarding medicine side effects among the general population: a cross-sectional survey. *Primary Health Care Research & Development*, 20(e153):1–8, 2019. ISSN 1463-4236. doi: 10.1017/S1463423619000574.

[226] K. Ogata, A. Kimura, N. Nakazawa, M. Suzuki, T. Yanoma, Y. Ubukata, K. Iwamatsu, N. Kogure, M. Yanai, and H. Kuwano. Long-Term Imatinib Treatment for Patients with Unresectable or Recurrent Gastrointestinal Stromal Tumors. *Digestion*, 97(1): 20–25, 2018. ISSN 14219867. doi: 10.1159/000484102.

[227] I. Omar and E. Harris. The Use of Social Media in ADR Monitoring and Reporting. *Journal of Pharmacovigilance*, 4(6), 2016. ISSN 2329-6887. doi: 10.4172/2329-6887.1000223.

[228] D. Pappa and L. K. Stergioulas. Harnessing social media data for pharmacovigilance: a review of current state of the art, challenges and future directions. *International Journal of Data Science and Analytics*, 8(2):113–135, 2019. ISSN 23644168. doi: 10.1007/s41060-019-00175-3.

[229] A. Park, A. L. Hartzler, J. Huh, D. W. Mcdonald, and W. Pratt. Automatically Detecting Failures in Natural Language Processing Tools for Online Community Text. *J Med Internet Res*, 17(8), 2015. doi: 10.2196/jmir.4612.

[230] J. Patrick, M. Sabbagh, S. Jain, and H. Zheng. Spelling correction in clinical notes with emphasis on first suggestion accuracy. In *2nd Workshop on Building and Evaluating Resources for Biomedical Text Mining*, pages 2–8, 2010.

[231] M. J. Paul and M. Dredze. A Model for Mining Public Health Topics from Twitter. Technical report, Johns Hopkins University, 2011. URL `http://www.michaeljpaul.com/files/2011.tech.twitter_health.pdf`.

[232] Y. Peng, S. Yan, and Z. Lu. Transfer Learning in Biomedical Natural Language Processing: An Evaluation of BERT and ELMo on Ten Benchmarking Datasets. In *Proceedings of the BioNLP 2019 workshop*, pages 58–65, 2019.

[233] L. Philips. The double metaphone search algorithm. *C/C++ Users J.*, 18(6):38–43, 2000. ISSN 1075-2838.

[234] C. M. Phillips, A. Parmar, H. Guo, D. Schwartz, W. Isaranuwatchai, J. Beca, W. Dai, J. Arias, S. Gavura, and K. K. Chan. Assessing the efficacy-effectiveness gap for cancer therapies: A comparison of overall survival and toxicity between clinical trial and population-based, real-world data for contemporary parenteral cancer therapeutics. *Cancer*, 126(8):1717–1726, 2020. ISSN 1097-0142. doi: 10.1002/CNCR.32697.

[235] P. Pimpalkhute, A. Patki, A. Nikfarjam, and G. Gonzalez. Phonetic spelling filter for keyword selection in drug mention mining from social media. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*, 2014:90–5, 2014. ISSN 2153-4063.

[236] K. Plueschke, P. McGettigan, A. Pacurariu, X. Kurz, and A. Cave. EU-funded initiatives for real world evidence: Descriptive analysis of their characteristics and relevance for regulatory decision-making. *BMJ Open*, 8(6):21864, 2018. ISSN 20446055. doi: 10.1136/bmjopen-2018-021864.

[237] G. R. Polich. Rare disease patient groups as clinical researchers. *Drug Discovery Today*, 17(3-4):167–172, 2012. ISSN 13596446. doi: 10.1016/j.drudis.2011.09.020.

[238] J. Pols. Knowing Patients: Turning Patient Knowledge into Science. *Science, Technology, & Human Values*, 39(1):73–97, 2014. doi: 10.1177/0162243913504306.

[239] H. Poort, W. T. van der Graaf, R. Tielen, M. Vlenterie, J. A. Custers, J. B. Prins, C. A. Verhagen, M. F. Gielissen, and H. Knoop. Prevalence, Impact, and Correlates of Severe Fatigue in Patients With Gastrointestinal Stromal Tumors. *Journal of pain and symptom management*, 52(2):265–271, 2016. ISSN 1873-6513. doi: 10.1016/J.JPAINSYMMAN.2016.02.019.

[240] J. Popay and G. Williams. Public health research and lay knowledge. *Social Science and Medicine*, 42(5):759–768, 1996. ISSN 02779536. doi: 10.1016/0277-9536(95)00341-X.

[241] S. Pradhan, N. Elhadad, W. Chapman, S. Manandhar, and G. Savova. SemEval-2014 Task 7: Analysis of Clinical Text. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 54–62, 2014. doi: 10.3115/v1/s14-2007.

[242] J. Price. What Can Big Data Offer the Pharmacovigilance of Orphan Drugs? *Clinical Therapeutics*, 38(12):2533–2545, 2016. ISSN 1879114X. doi: 10.1016/j.clinthera.2016.11.009.

[243] L. Prior. Belief, knowledge and expertise: the emergence of the lay expert in medical sociology. *Sociology of Health & Illness*, 25(3):41–57, 2003. ISSN 01419889. doi: 10.1111/1467-9566.00339.

[244] C. A. Radawski, T. A. Hammad, S. Colilla, P. Coplan, K. Hornbuckle, E. Freeman, M. Y. Smith, R. E. Sobel, P. Bahri, A. E. Arias, and D. Bennett. The utility of real-world evidence for benefit-risk assessment, communication, and evaluation of pharmaceuticals: Case studies. *Pharmacoepidemiology and Drug Safety*, 29(12):1532–1539, 2020. ISSN 1053-8569. doi: 10.1002/pds.5167.

[245] L. A. Ramshaw and M. P. Marcus. Text chunking using transformation-based learning. In S. Armstrong, K. Church, P. Isabelle, S. Manzi, E. Tzoukermann, and D. Yarowsky, editors, *Natural Language Processing Using Very Large Corpora*, pages 157–176, Dordrecht, 1999. Springer Netherlands. ISBN 978-94-017-2390-9. doi: 10.1007/978-94-017-2390-9_10.

[246] P. Reichardt. The Story of Imatinib in GIST-a Journey through the Development of a Targeted Therapy. *Oncol Res Treat*, 41:472–477, 2018. doi: 10.1159/000487511. URL www.karger.com/ort.

[247] N. Reimers and I. Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, pages 3982–3992, 2019.

[248] M. Reynaert. *Text-Induced Spelling Correction*. PhD thesis, Tilburg University, 2005.

[249] T. Richter, S. Nestler-Parr, R. Babela, Z. M. Khan, T. Tesoro, E. Molsen, and D. A. Hughes. Rare Disease Terminology and Definitions-A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. *Value in Health*, 18(6): 906–914, 2015. ISSN 15244733. doi: 10.1016/j.jval.2015.05.008.

[250] A. Rios and R. Kavuluru. Few-Shot and Zero-Shot Multi-Label Learning for Structured Label Spaces. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*, pages 3132–3142, 2018. doi: 10.18653/V1/D18-1352.

[251] K. M. Robinson. Unsolicited narratives from the internet: A rich source of qualitative data. *Qualitative Health Research*, 11(5):706–714, 2001. ISSN 10497323. doi: 10.1177/104973201129119398.

[252] L. Rolfes, F. van Hunsel, K. Taxis, and E. van Puijenbroek. The Impact of Experiencing Adverse Drug Reactions on the Patient's Quality of Life: A Retrospective Cross-Sectional Study in the Netherlands. *Drug Safety*, 39(8):769–776, 2016. ISSN 11791942. doi: 10.1007/s40264-016-0422-0.

[253] L. Rolfes, F. van Hunsel, L. van der Linden, K. Taxis, and E. van Puijenbroek. The Quality of Clinical Information in Adverse Drug Reaction Reports by Patients and Healthcare Professionals: A Retrospective Comparative Analysis. *Drug Safety*, 40(7): 607–614, 2017. ISSN 11791942. doi: 10.1007/s40264-017-0530-5.

[254] S. Rosenthal and K. McKeown. I couldn't agree more: The role of conversational structure in agreement and disagreement detection in online discussions. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 168–177, Prague, Czech Republic, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4625.

[255] P. Ruch, R. Baud, and A. Geissbühler. Using lexical disambiguation and named-entity recognition to improve spelling correction in the electronic patient record. *Artificial Intelligence in Medicine*, 29(1):169 – 184, 2003. ISSN 0933-3657. doi: https://doi.org/10.1016/S0933-3657(03)00052-6.

[256] M. Ryu and K. Lee. Knowledge Distillation for BERT Unsupervised Domain Adaptation. *arXiv*, 2020. URL https://arxiv.org/abs/2010.11478v2.

[257] A. Sakhovskiy, Z. Miftahutdinov, and E. Tutubalina. KFU NLP Team at SMM4H 2021 Tasks: Cross-lingual and Cross-modal BERT-based Models for Adverse Drug Effects. In *Proceedings of the Sixth Social Media Mining for Health Workshop 2021*, pages 39–43, 2021. doi: 10.18653/v1/2021.smm4h-1.6.

[258] H. Sampathkumar, X.-W. Chen, and B. Luo. Mining Adverse Drug Reactions from online healthcare forums using Hidden Markov Model. *BMC Medical Informatics and Decision Making*, 14, 2014.

[259] O. Sangha, G. Stucki, M. H. Liang, A. H. Fossel, and J. N. Katz. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis and rheumatism*, 49(2):156–163, 2003. ISSN 0004-3591. doi: 10.1002/ART.10993.

[260] V. Sanh, L. Debut, J. Chaumond, and T. Wolf. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. In *5th Workshop on Energy Efficient Machine Learning and Cognitive Computing - NeurIPS 2019*, 2019.

[261] A. Sarker. A customizable pipeline for social media text normalization. *Social Network Analysis and Mining*, 7(1):45, 2017. ISSN 1869-5450. doi: 10.1007/s13278-017-0464-z.

[262] A. Sarker and G. Gonzalez. Portable automatic text classification for adverse drug reaction detection via multi-corpus training. *Journal of Biomedical Informatics*, 53:196–207, 2015. ISSN 1532-0464. doi: 10.1016/J.JBI.2014.11.002.

[263] A. Sarker and G. Gonzalez. A corpus for mining drug-related knowledge from Twitter chatter: Language models and their utilities. *Data in Brief*, 10:122–131, 2017. ISSN 2352-3409. doi: 10.1016/J.DIB.2016.11.056.

[264] A. Sarker and G. Gonzalez. HLP@UPenn at SemEval-2017 Task 4A: A simple, self-optimizing text classification system combining dense and sparse vectors. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pages 640–643, Vancouver, Canada, 2017. Association for Computational Linguistics. doi: 10.18653/v1/S17-2105.

[265] A. Sarker and G. Gonzalez-Hernandez. Overview of the Second Social Media Mining for Health (SMM4H) Shared Tasks at AMIA 2017. In *Proceedings of the Second Workshop on Social Media Mining for Health Applications (SMM4H)*, 2017.

[266] A. Sarker, R. Ginn, A. Nikfarjam, K. O'Connor, K. Smith, S. Jayaraman, T. Upadhaya, and G. Gonzalez. Utilizing social media data for pharmacovigilance: A review. *Journal of Biomedical Informatics*, 54:202–212, 2015. ISSN 1532-0464. doi: 10.1016/J.JBI.2015.02.004.

[267] A. Sarker, K. O'Connor, R. Ginn, M. Scotch, K. Smith, D. Malone, and G. Gonzalez. Social Media Mining for Toxicovigilance: Automatic Monitoring of Prescription Medication Abuse from Twitter. *Drug Safety*, 39(3):231–240, 2016. ISSN 0114-5916. doi: 10.1007/s40264-015-0379-4.

[268] A. Sarker, M. Belousov, J. Friedrichs, K. Hakala, S. Kiritchenko, F. Mehryary, S. Han, T. Tran, A. Rios, R. Kavuluru, B. de Bruijn, F. Ginter, D. Mahata, S. M. Mohammad, G. Nenadic, and G. Gonzalez-Hernandez. Data and systems for medication-related text classification and concept normalization from Twitter: insights from the Social Media Mining for Health (SMM4H)-2017 shared task. *Journal of the American Medical Informatics Association*, 25(10):1274–1283, 2018. ISSN 1067-5027. doi: 10.1093/jamia/ocy114.

[269] A. Sarker, S. Lakamana, W. Hogg-Bremer, A. Xie, M. A. Al-Garadi, and Y.-C. Yang. Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource. *Journal of the American Medical Informatics Association*, 27(8):1310–1315, 2020. ISSN 1527-974X. doi: 10.1093/jamia/ocaa116.

[270] C. Seale, J. Charteris-Black, A. MacFarlane, and A. McPherson. Interviews and internet forums: a comparison of two sources of qualitative data. *Qualitative health research*, 20(5):595–606, 2010. ISSN 1049-7323. doi: 10.1177/1049732309354094.

[271] I. Segura-Bedmar, P. Martínez, R. Revert, and J. Moreno-Schneider. Exploring Spanish health social media for detecting drug effects. *BMC Medical Informatics and Decision Making*, 15(2):1–9, 2015. ISSN 14726947. doi: 10.1186/1472-6947-15-S2-S6.

[272] J. Shang, L. Liu, X. Gu, X. Ren, T. Ren, and J. Han. Learning named entity tagger using domain-specific dictionary. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*, pages 2054–2064, 2020. doi: 10.18653/v1/d18-1230.

[273] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. ISSN 0005-8580. doi: 10.1002/j.1538-7305.1948.tb01338.x.

[274] P. Shenoy and A. Harugeri. Elderly patients' participation in clinical trials. *Perspectives in Clinical Research*, 6(4):184, 2015. ISSN 2229-3485. doi: 10.4103/2229-3485.167099.

[275] Y. Si, J. Wang, H. Xu, and K. Roberts. Enhancing clinical concept extraction with contextual embeddings. *Journal of the American Medical Informatics Association*, 26(11):1297–1304, 2019. doi: https://doi.org/10.1093/jamia/ocz096.

[276] R. Sloane, O. Osanlou, D. Lewis, D. Bollegala, S. Maskell, and M. Pirmohamed. Social media and pharmacovigilance: A review of the opportunities and challenges. *British journal of clinical pharmacology*, 80(4):910–20, 2015. ISSN 1365-2125. doi: 10.1111/bcp.12717.

[277] E. Smailhodzic, W. Hooijsma, A. Boonstra, and D. J. Langley. Social media use in healthcare: A systematic review of effects on patients and on their relationship with healthcare professionals. *BMC Health Services Research*, 16(1), 2016. ISSN 1472-6963. doi: 10.1186/s12913-016-1691-0.

[278] M. V. Smeden, C. Moons, I. Kant, and H. V. Os. Leidraad voor kwalitatieve diagnostische en prognostische toepassingen van AI in de zorg. Technical report, Ministerie van Volksgezondheid, Welzijn en Sport, 2021. URL https://www.datavoorgezondheid.nl/documenten/publicaties/2021/12/17/leidraad-kwaliteit-ai-in-de-zorg.

[279] D. Smedt, R. H. . Jaarsma, T. . Ranchor, A. V. . Van Der Meer, K. . Groenier, K. H. . Haaijer-Ruskamp, and F. M. . Denig. Coping with adverse drug events in patients with heart failure: Exploring the role of medication beliefs and perceptions. *Psychology & Health*, 27(5):570–587, 2012. doi: 10.1080/08870446.2011.605886.

[280] K. Smith, S. Golder, A. Sarker, Y. Loke, K. O'Connor, and G. Gonzalez-Hernandez. Methods to Compare Adverse Events in Twitter to FAERS, Drug Information Databases, and Systematic Reviews: Proof of Concept with Adalimumab. *Drug Safety*, 41(12):1397–1410, 2018. ISSN 11791942. doi: 10.1007/s40264-018-0707-6.

[281] L. N. Smith. A disciplined approach to neural network hyper-parameters: Part 1 - learning rate, batch size, momentum, and weight decay. *ArXiv*, 2018. URL https://arxiv.org/pdf/1803.09820.pdf.

[282] S. C. Sodergren, A. White, F. Efficace, M. Sprangers, D. Fitzsimmons, A. Bottomley, and C. D. Johnson. Systematic review of the side effects associated with tyrosine kinase inhibitors used in the treatment of gastrointestinal stromal tumours on behalf of the EORTC Quality of Life Group. *Critical reviews in oncology/hematology*, 91(1):35–46, 2014. ISSN 1879-0461. doi: 10.1016/J.CRITREVONC.2014.01.002.

[283] S. C. Sodergren, S. J. Wheelwright, D. Fitzsimmons, F. Efficace, M. Sprangers, P. Fayers, A. Harle, H. Schmidt, A. Bottomley, A. S. Darlington, C. Benson, A. Bredart, L. Hentschel, J. I. Arraras, G. Ioannidis, M. Leahy, I. Lugowska, O. Nicolatou-Galitis, D. Petranovic, G. E. Rohde, V. Vassiliou, and C. D. Johnson. Developing Symptom Lists for People with Cancer Treated with Targeted Therapies. *Targeted oncology*, 16 (1):95–107, 2021. ISSN 1776-260X. doi: 10.1007/S11523-020-00769-Z.

[284] C. Song, S. Zhang, N. Sadoughi, P. Xie, and E. Xing. Generalized Zero-Shot Text Classification for ICD Coding. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20) Code*, pages 4018–4024, 2020.

[285] K. Søreide, O. M. Sandvik, J. A. Søreide, V. Giljaca, A. Jureckova, and V. R. Bulusu. Global epidemiology of gastrointestinal stromal tumours (GIST): A systematic review of population-based cohort studies. *Cancer Epidemiology*, 40:39–46, 2016. ISSN 1877783X. doi: 10.1016/j.canep.2015.10.031.

[286] M. H. Stanfill, M. Williams, S. H. Fenton, R. A. Jenders, and W. R. Hersh. A systematic literature review of automated clinical coding and classification systems. *JAMIA*, 17: 646–651, 2010. doi: 10.1136/jamia.2009.001024.

[287] S. Stieglitz, M. Mirbabaie, B. Ross, and C. Neuberger. Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International*

Journal of Information Management, 39:156–168, 2018. ISSN 02684012. doi: 10.1016/j.ijinfomgt.2017.12.002.

[288] A. Stolcke, K. Ries, N. Coccaro, E. Shriberg, R. Bates, D. Jurafsky, P. Taylor, R. Martin, C. Van Ess-Dykema, and M. Meteer. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26(3):339–374, 2000.

[289] B. H. Stricker and B. M. Psaty. Detection, verification, and quantification of adverse drug reactions. *BMJ*, 329:44–47, 2004. doi: 10.1136/bmj.329.7456.44.

[290] L. Sun, K. Hashimoto, W. Yin, A. Asai, J. Li, P. Yu, and C. Xiong. Adv-BERT: BERT is not robust on misspellings! Generating nature adversarial samples on BERT. *arXiv*, 2020. ISSN 23318422. URL `https://arxiv.org/abs/2003.04985`.

[291] M. Sung, H. Jeon, J. Lee, and J. Kang. Biomedical entity representations with synonym marginalization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3641–3650, 2020. doi: 10.18653/v1/2020.acl-main.335.

[292] D. Supranovich and V. Patsepnia. Ihs_rd: Lexical normalization for english tweets. In *Proceedings of the Workshop on Noisy User-generated Text*, pages 78–81, Beijing, China, 2015. Association for Computational Linguistics. doi: 10.18653/v1/W15-4311.

[293] D. Sánchez, M. Batet, and A. Viejo. Utility-preserving privacy protection of textual healthcare documents. *Journal of Biomedical Informatics*, 52:189 – 198, 2014. doi: https://doi.org/10.1016/j.jbi.2014.06.008.

[294] H. Taipale, J. Schneider-Thoma, J. Pinzón-Espinosa, J. Radua, O. Efthimiou, C. H. Vinkers, E. Mittendorfer-Rutz, N. Cardoner, L. Pintor, A. Tanskanen, A. Tomlinson, P. Fusar-Poli, A. Cipriani, E. Vieta, S. Leucht, J. Tiihonen, and J. J. Luykx. Representation and Outcomes of Individuals With Schizophrenia Seen in Everyday Practice Who Are Ineligible for Randomized Clinical Trials. *JAMA Psychiatry*, 2022. ISSN 2168-622X. doi: 10.1001/JAMAPSYCHIATRY.2021.3990.

[295] B. Tang, Q. Chen, X. Wang, Y. Wu, Y. Zhang, M. Jiang, J. W. Wang, and H. Xu. Recognizing Disjoint Clinical Concepts in Clinical Text Using Machine Learning-based Methods. In *AMIA Annu Symp Proc.*, pages 1184–1193, 2015.

[296] B. Tang, J. Hu, X. Wang, and Q. Chen. Recognizing Continuous and Discontinuous Adverse Drug Reaction Mentions from Social Media Using LSTM-CRF. *Wireless Communications and Mobile Computing*, 2018. doi: 10.1155/2018/2379208.

[297] Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29 (1):24–54, 2010. ISSN 0261927X. doi: 10.1177/0261927X09351676.

[298] E. Tjong Kim Sang and F. de Meulder. Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition. In *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, pages 142–147, 2003.

[299] G. Tortoreto, E. A. Stepanov, A. Cervone, M. Dubiel, and G. Riccardi. Affective Behaviour Analysis of On-line User Interactions: Are On-line Support Groups more Therapeutic than Twitter? In *Proceedings of the 4th Social Media Mining for Health Applications (#SMM4H) Workshop & Shared Task*, pages 79–88, 2019. doi: 10.18653/v1/W19-3211.

[300] H. Tu, Z. Ma, A. Sun, and X. Wang. When MetaMap Meets Social Media in Healthcare: Are the Word Labels Correct? In *Asia Information Retrieval Symposium*, pages 356–362. Springer, 2016. doi: 10.1007/978-3-319-48051-0_31.

[301] Z. Tufekci. Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. In *Proceedings of the eighth international AAAI conference on weblogs and social media.*, pages 505–514, 2014. doi: 10.1016/0022-5193(78)90170-4.

[302] E. Tutubalina, Z. Miftahutdinov, S. Nikolenko, and V. Malykh. Medical concept normalization in social media posts with recurrent neural networks. *Journal of Biomedical Informatics*, 84:93–102, 2018. ISSN 1532-0464. doi: 10.1016/J.JBI.2018.06.006.

[303] E. Tutubalina, I. Alimova, Z. Miftahutdinov, A. Sakhovskiy, V. Malykh, and S. Nikolenko. The Russian Drug Reaction Corpus and Neural Models for Drug Reactions and Effectiveness Detection in User Reviews. *Bioinformatics*, 37(2), 2020. doi: 10.1093/bioinformatics/btaa675.

[304] E. Tutubalina, A. Kadurin, and Z. Miftahutdinov. Fair Evaluation in Concept Normalization: a Large-scale Comparative Analysis for BERT-based Models. In *COLING 2020*, 2020. doi: 10.18653/v1/2020.coling-main.588.

[305] U.S Congress Office of Technology Assessment. *Pharmaceutical R&D: Costs, Risks, and Rewards*. U.S. Government Printing Office, Washington, DC, 1993. ISBN 0-16-041658-2. URL `https://ota.fas.org/reports/9336.pdf`.

[306] US Department of Health & Human Services. FAQs About Rare Diseases, 2021. URL `https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases`.

[307] US Department of Health and Human Services. Common Terminology Crtieria for Adverse Drug Events (CTCAE) version 5. Technical report, US Department of Health and Human Services, 2017. URL `https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcae_v5_quick_reference_5x7.pdf`.

[308] U.S. Food and Drug Administration (FDA). GLEEVEC (imatinib mesylate) tablets label. Technical report, U.S. Food and Drug Administration (FDA),

2008. URL https://www.accessdata.fda.gov/drugsatfda_docs/label/2008/021588s024lbl.pdf.

[309] U.S. Food and Drug Administration (FDA). SUTENT (sunitinib malate) capsules label. Technical report, U.S. Food and Drug Administration (FDA), 2011. URL https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/021938s13s17s18lbl.pdf.

[310] U.S. Food and Drug Administration (FDA). STIVARGA (regorafenib) tablets label. Technical report, U.S. Food and Drug Administration (FDA), 2017. URL https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/203085s007lbl.pdf.

[311] U.S. Food and Drug Administration (FDA). FDA approves ripretinib for advanced gastrointestinal stromal tumor, 2020. URL https://www.fda.gov/drugs/drug-approvals-and-databases/fda-approves-ripretinib-advanced-gastrointestinal-stromal-tumor.

[312] U.S. Food and Drug Administration (FDA). Grapefruit juice can affect how well some medicines work, 2021. URL https://www.fda.gov/consumers/consumer-updates/grapefruit-juice-and-some-drugs-dont-mix.

[313] U.S. National Library of Medicine. Unified Medical Language System (UMLS). URL https://www.nlm.nih.gov/research/umls/.

[314] U.S. National Library of Medicine. RxNorm, 2020. URL https://www.nlm.nih.gov/research/umls/rxnorm/.

[315] Ö. Uzuner, B. R. South, S. Shen, and S. L. DuVall. 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association*, 18(5):552–556, 2011. ISSN 10675027. doi: 10.1136/amiajnl-2011-000203.

[316] S. van Buuren and K. Groothuis-Oudshoorn. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3):1–67, 2011. ISSN 1548-7660. doi: 10.18637/JSS.V045.I03.

[317] L. V. Van De Poll-Franse, N. Horevoorts, M. V. Eenbergen, J. Denollet, J. A. Roukema, N. K. Aaronson, A. Vingerhoets, J. W. Coebergh, J. De Vries, M. L. Essink-Bot, and F. Mols. The Patient Reported Outcomes Following Initial treatment and Long term Evaluation of Survivorship registry: Scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. *European Journal of Cancer*, 47(14):2188–2194, 2011. ISSN 09598049. doi: 10.1016/j.ejca.2011.04.034.

[318] R. van der Goot and G. van Noord. Monoise: Modeling noise using a modular normalization system. *Computational Linguistics in the Netherlands Journal*, 7:129–144, 2017.

[319] M. C. van Eenbergen, L. V. van de Poll-Franse, P. Heine, and F. Mols. The Impact of Participation in Online Cancer Communities on Patient Reported Outcomes: Systematic Review. *JMIR Cancer*, 3(2), 2017. doi: 10.2196/cancer.7312.

[320] F. Van Hunsel, Susan De Waal, and L. Härmark. The contribution of direct patient reported ADRs to drug safety signals in the Netherlands from 2010 to 2015. *Pharmacoepidemiol Drug Saf*, 26:977–983, 2017. doi: 10.1002/pds.4236.

[321] J. van Stekelenborg, J. Ellenius, S. Maskell, T. Bergvall, O. Caster, N. Dasgupta, J. Dietrich, S. Gama, D. Lewis, V. Newbould, S. Brosch, C. E. Pierce, G. Powell, A. Ptaszyńska-Neophytou, A. F. Wiśniewski, P. Tregunno, G. N. Norén, and M. Pirmohamed. Recommendations for the Use of Social Media in Pharmacovigilance: Lessons from IMI WEB-RADR. *Drug Safety*, 42(12):1393–1407, 2019. ISSN 11791942. doi: 10.1007/s40264-019-00858-7.

[322] C. F. van Uden-Kraan, C. H. Drossaert, E. Taal, C. E. Lebrun, K. W. Drossaers-Bakker, W. M. Smit, E. R. Seydel, and M. A. van de Laar. Coping with somatic illnesses in online support groups: Do the feared disadvantages actually occur? *Computers in Human Behavior*, 24(2):309–324, 2008. ISSN 07475632. doi: 10.1016/j.chb.2007.01.014.

[323] C. F. van Uden-Kraan, C. H. Drossaert, E. Taal, E. R. Seydel, and M. A. van de Laar. Self-reported differences in empowerment between lurkers and posters in online patient support groups. *Journal of medical Internet research*, 10(2):1–9, 2008. ISSN 14388871. doi: 10.2196/jmir.992.

[324] C. F. van Uden-Kraan, C. H. Drossaert, E. Taal, B. R. Shaw, E. R. Seydel, and M. A. F. J. van de Laar. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. *Qualitative Health Research*, 18(3):405–417, 2008. ISSN 1049-7323. doi: 10.1177/1049732307313429.

[325] C. F. van Uden-Kraan, C. H. Drossaert, E. Taal, E. R. Seydel, and M. A. van de Laar. Participation in online patient support groups endorses patients' empowerment. *Patient Education and Counseling*, 74(1):61–69, 2009. ISSN 07383991. doi: 10.1016/j.pec.2008.07.044.

[326] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention Is All You Need. In *31st Conference on Neural Information Processing Systems (NIPS 2017)*, pages 5998–6008, 2017.

[327] S. Verberne. Context-sensitive spell checking based on word trigram probabilities. Master's thesis, Radboud University, 2002.

[328] S. Verberne, A. Batenburg, R. Sanders, and M. V. Eenbergen. Social processes of online empowerment on a cancer patient discussion form: using text mining to analyze linguistic patterns of empowerment processes. *JMIR Cancer*, 1(5), 2018. doi: 10.2196/cancer.9887.

[329] S. Verberne, A. Batenburg, R. Sanders, M. van Eenbergen, E. Das, and M. S. Lambooij. Analyzing empowerment processes among cancer patients in an online community: A text mining approach. *JMIR Cancer*, 5(1):e9887, 2019. ISSN 2369-1999. doi: 10.2196/cancer.9887.

[330] K. Verspoor, J. Cohn, S. Mniszewski, and C. Joslyn. A categorization approach to automated ontological function annotation. *Protein Science*, 15(6):1544–1549, 2006. ISSN 09618368. doi: 10.1110/ps.062184006.

[331] J. Verweij, P. G. Casali, J. Zalcberg, A. LeCesne, P. Reichardt, J.-Y. Blay, R. Issels, A. van Oosterom, P. C. Hogendoorn, M. Van Glabbeke, R. Bertulli, and I. Judson. Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial. *The Lancet*, 364(9440):1127–1134, 9 2004. ISSN 0140-6736. doi: 10.1016/S0140-6736(04)17098-0. URL https://www.sciencedirect.com/science/article/pii/S0140673604170980?via%3Dihub.

[332] P. Vijayaraghavan and D. Roy. Modeling human motives and emotions from personal narratives using external knowledge and entity tracking. In *Proceedings of the Web Conference 2021 (WWW '21)*, pages 529–540. ACM, 2021. ISBN 978-1-4503-8312-7. doi: 10.1145/3442381.3449997.

[333] J. E. Ware and C. D. Sherbourne. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical care*, 30(6):473–483, 1992. ISSN 0025-7079. doi: 10.1097/00005650-199206000-00002.

[334] WEB-RADR 2 project, 2021. URL https://www.snomed.org/news-and-events/articles/new-collaboration-SNOMED-ICH-MedDRA.

[335] D. Weissenbacher, A. Sarker, M. Paul, and G. Gonzalez-Hernandez. Overview of the Third Social Media Mining for Health (SMM4H) Shared Tasks at EMNLP 2018. In *Proceedings of the 3rd Social Media Mining for Health Applications (SMM4H) Workshop & Shared Task*, pages 13–16, 2018. doi: 10.18653/v1/W18-5904.

[336] D. Weissenbacher, A. Sarker, A. Klein, K. O'connor, A. Magge, and G. Gonzalez-Hernandez. Deep neural networks ensemble for detecting medication mentions in tweets. *Journal of the American Medical Informatics Association*, 26(12):1618–1626, 2019. ISSN 1527974X. doi: 10.1093/jamia/ocz156.

[337] D. Weissenbacher, A. Sarker, A. Magge, A. Daughton, K. O'Connor, M. Paul, and G. Gonzalez-Hernandez. Overview of the Fourth Social Media Mining for Health (#SMM4H) Shared Task at ACL 2019. In *Proceedings of the 4th Social Media Mining for Health Applications (#SMM4H) Workshop & Shared Task*, pages 21–30, 2019. doi: 10.18653/v1/W19-3203.

[338] D. William and D. Suhartono. Text-based Depression Detection on Social Media Posts: A Systematic Literature Review. *Procedia Computer Science*, 179:582–589, jan 2021. ISSN 1877-0509. doi: 10.1016/J.PROCS.2021.01.043.

[339] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, and J. Brew. HuggingFace's Transformers: State-of-the-art Natural Language Processing. *ArXiv*, 2019. URL https://arxiv.org/abs/1910.03771.

[340] World Health Organisation. The Safety of Medicines in Public Health Programmes: Pharmacovigilance, an essential tool. Technical report, World Health Organisation, 2006. URL https://apps.who.int/iris/handle/10665/43384.

[341] World Health Organization. The Importance of Pharmacovigilance. Technical report, World Health Organization, Geneva, 2002. URL https://apps.who.int/iris/handle/10665/42493.

[342] World Health Organization. Safety of Medicines: A guide to detecting and reporting adverse drug reactions, why health professionals need to take action. Technical report, World Health Organization, Geneva, 2002. URL https://apps.who.int/iris/handle/10665/67378.

[343] Y. Wu, B. Tang, M. Jiang, S. Moon, J. C. Denny, and H. Xu. Clinical acronym/abbreviation normalization using a hybrid approach. In *CLEF (Working Notes)*, 2013.

[344] C. C. Yang, H. Yang, L. Jiang, and M. Zhang. Social media mining for drug safety signal detection. In *Proceedings of the 2012 international workshop on Smart health and wellbeing - SHB '12*, pages 33–40, 2012. ISBN 9781450317122. doi: 10.1145/2389707.2389714.

[345] A. Yates and N. Goharian. ADRTrace: Detecting Expected and Unexpected Adverse Drug Reactions from User Reviews on Social Media Sites. In *ECIR 2013: Advances in Information Retrieval*, pages 816–819, 2013. doi: 10.1007/978-3-642-36973-5_92.

[346] S. Yeleswarapu, A. Rao, T. Joseph, V. Govindakrishnan Saipradeep, and R. Srinivasan. A pipeline to extract drug-adverse event pairs from multiple data sources. *BMC medical informatics and decision making*, 14(13), 2014. doi: 10.1186/1472-6947-14-13.

[347] Y. Zang, B. Hou, F. Qi, Z. Liu, X. Meng, and M. Sun. Learning to attack: Towards textual adversarial attacking in real-world situations. *arXiv*, 2020. ISSN 23318422. URL https://arxiv.org/abs/2009.09192.

[348] Q. Zeng and T. Tse. Exploring and developing consuming health vocabulary. *J Am Med Inform Assoc*, 13(1):24–29, 2006. doi: 10.1197/jamia.M1761.A.

[349] B. Zhang, X. Zhang, Y. Liu, L. Cheng, and Z. Li. Matching Distributions between Model and Data: Cross-domain Knowledge Distillation for Unsupervised Domain Adaptation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, pages 5423–5433, 2021.

[350] M. Zhang, M. Zhang, C. Ge, Q. Liu, J. Wang, J. Wei, and K. Q. Zhu. Automatic discovery of adverse reactions through Chinese social media. *Data Mining and Knowledge Discovery*, 33(4):848–870, 2019. ISSN 1573756X. doi: 10.1007/s10618-018-00610-2.

[351] Y. Zhang, J. Wang, B. Tang, Y. Wu, M. Jiang, Y. Chen, and H. Xu. UTH_CCB: A report for SemEval 2014 – Task 7 Analysis of Clinical Text. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 802–806, 2014. doi: 10.3115/v1/s14-2142.

[352] X. Zhou, A. Zheng, J. Yin, R. Chen, X. Zhao, W. Xu, W. Cheng, T. Xia, and S. Lin. Context-Sensitive Spelling Correction of Consumer-Generated Content on Health Care. *JMIR Medical Informatics*, 3(3):e27, 2015. ISSN 2291-9694. doi: 10.2196/medinform.4211.

[353] M. Zolnoori, K. W. Fung, T. B. Patrick, P. Fontelo, H. Kharrazi, A. Faiola, N. D. Shah, Y. S. Shirley Wu, C. E. Eldredge, J. Luo, M. Conway, J. Zhu, S. K. Park, K. Xu, and H. Moayyed. The PsyTAR dataset: From patients generated narratives to a corpus of adverse drug events and effectiveness of psychiatric medications. *Data in brief*, 24, 2019. ISSN 2352-3409. doi: 10.1016/j.dib.2019.103838.

[354] A. Zubiaga, E. Kochkina, M. Liakata, R. Procter, M. Lukasik, K. Bontcheva, T. Cohn, and I. Augenstein. Discourse-Aware Rumour Stance Classification in Social Media Using Sequential Classifiers. *Information Processing & Management*, 54(2):273–390, 2018. doi: 10.1016/j.ipm.2017.11.009.