REFERENCES

[1] A. Alaphilippe, A. Gizikis, C. Hanot, “Automated tackling of disinformation”, 2019, [Online]. Available: https://www.europarl.europa.eu/RegData/etudes

[2] General Data Protection Regulation [Online]. Available: https://en.wikipedia.org/wiki/General_Data_Protection_Regulation

[3] The “localisation” of Russian citizens’ personal data [Online]. Available: https://home.kpmg/be/en/home/insights/2018/09/the-localisation-of-russian-citizens-personal-data.html

[4] B. B. Mehta, U. P. Rao, “Privacy preserving unstructured big data analytics – issues and challenges”, in Proc. International Conference on Security and Privacy, Nagpur, India, 2015, pp. 120-124.

[5] R. Mendes, J. P. Vilela, “Privacy-preserving data mining: Methods, metrics, and applications,” in IEEE Access, 2017, pp. 10562–10582.

[6] P. Usha, R. Shriram, S. Sathishkumar, “Multiple sensitive attributes-based privacy preserving data mining using k-anonymity” in Int. J. Sci. Eng. Res. 5(4), 2014

[7] A. Kaur, “A hybrid approach of privacy preserving data mining using suppression and perturbation techniques,” in Proc. IEEE International Conference on Innovative Mechanisms for Industry Applications, ICIMIA 2017, India, 2017, pp. 306–311

[8] D. Kumari, Y. Vineela, T. Krishna, B. Kumar, “Analyzing and performing Privacy Preserving Data Mining on medical databases” in Indian Journal of Science and Technology. 2016 May; pp. 1–9.

[9] L. Sweeney, “Achieving k-Anonymity Privacy Protection Using Generalization and Suppression” in International Journal on Uncertainty, Fuzziness and Knowledge based Systems 10(5), 2002, pp. 571–588

[10] L. Zhang, W. Zhang, “Generalization-based privacy-preserving data collection”, in: Proceedings of the 10th International Conference on Data Warehousing and Knowledge Discovery, DaWak, 2008.

[11] S. Hajian, J. Domingo-Ferrer, O. Farr`as, “Generalization-based privacy preservation and discrimination prevention in data publishing and mining” DMKD, 2014, pp. 1158–1188

[12] W. Yu, P. Lv, N. Chen, "Multi-Attribute Generalization Method in Privacy Preserving Data Publishing," in 2010 2nd International Conference on E-business and Information System Security, Wuhan, 2010, pp. 1-4.

[13] W.K. Wong, N. Mamoulis, D.W.L. Cheung, “Non-homogeneous generalization in privacy preserving data publishing” In SIGMOD, 2010, pp. 747–758

[14] A. S. M. Hasan, Q. Jiang, J. Luo, C. Li, L. Chen, “An effective value swapping method for privacy preserving data publishing” in Security and Communication Networks 9, 3219, 2016

[15] A. Richard, “Controlled Data-Swapping Techniques for Masking Public Use Micro Datasets”

[16] S.E. Fienberg, and J. McIntyre, “Data Swapping: Variations on a Theme by Dalenius and Reiss”, in Journal of Official Statistics, 9, pp. 383-406.

[17] V.S. Susan, T. Christopher, “Anatomisation with slicing: a new privacy preservation approach for multiple sensitive attributes” Springerplus 5(1), 964, 2016

[18] T. G. Marathe, M. N. Raverkar, S.G. Suryawanshi, M.F. More, “Preserving the Privacy of User by Using Anonymization Techniques” in International Journal of Sustainable Development Research, 2017

[19] R. Oksvort, “A Prototype for Learning Privacy-Preserving Data Publishing”, 2017

[20] B. Thuraisingham, M. Kantarcioglu, L. Liu, “Perturbation based privacy preserving data mining techniques for real-world data”, Doctoral Dissertation, 2008

[21] N. Patel, S. Patel, “A Study on Data Perturbation Techniques in Privacy Preserving Data Mining” in International Research Journal of Engineering and Technology, 2015

[22] B. C. Chen, D. Kifer, K. LeFevre, A. Machanavajjhala , “Privacy-preserving data publishing,” in Proc. Foundations and Trends in Databases Conference, 2009, pp. 1 – 167.

[23] P. Samarati, L. Sweeney, “Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and cell suppression”, Technical report, SRI International, 1998.

[24] G. Navarro-Arribas, V. Torra, A. Erola, J. Castellà-Roca, “User k-anonymity for privacy preserving data mining of query logs”, Information Processing and Management 48 (3), 2012, pp. 476–487.

[25] S. Ni, M. Xie, M, Q. Qian, “Clustering Based K-anonymity Algorithm for Privacy Preservation” in International Journal of Network Security, 2017, pp. 1062–1071.

[26] A. Machanavajjhala, J. Gehrke, D. Kifer,, M. Venkitasubramaniam., “l-diversity: Privacy beyond k-anonymity”, in Proc. 22nd International Conference on Data Engineering (ICDE). IEEE Computer Society, 2006

[27] N. li, T. Li, S. Venkatasubramanian, “t-Closeness: Privacy Beyond k-Anonymity and l-Diversity”, in IEEE 23rd International Conference on Data Engineering, 2007

[28] Z. Huang, “Privacy-Preserving Algorithms for Genomic Data”

[29] C. Uhlerop, A. Slavković, S. E. Fienberg, “Privacy-Preserving Data Sharing for Genome-Wide Association Studies”, 2015

[30] J. Gardner and L. Xiong, “An integrated framework for de-identifying heterogeneous data”, in Proc. Data and Knowledge Engineering, 2009, pp. 1441-1451

[31] K. Liu, K. Das, T. Grandison, and H. Kargupta. “Privacy-preserving data analysis on graphs and social networks”, In H. Kargupta, J. Han, P. Yu, R. Motwani, and V. Kumar, editors, Next Generation Data Mining. CRC Press, 2008

[32] J. Vadisala and V. K. Vatsavayi, “Challenges in Social Network Data Privacy” in International Journal of Computational Intelligence Research (IJCIR), vol -13, 2017, pp. 965-979

[33] C. Liu, P. Mittal. “Linkmirage: Enabling privacy-preserving analytics on social relationships”, in NDSS, 2016, pp. 21-24

[34] A. Meyerson and R. Williams, “On the complexity of optimal k-anonymity,” in Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS), 2004.

[35] V. S. Iyengar, “Transforming data to satisfy privacy constraints,” in ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2002.

[36] R. Bayardo and R. Agrawal, “Data privacy through optimal k-anonymity,” in Proceedings of the 21st International Conference on Data Engineering (ICDE), 2005.

[37] K. LeFevre, D. DeWitt, and R. Ramakrishnan, “Workload-Aware Anonymization,” in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2006.

[38] X. Xiao and Y. Tao. “Personalized privacy preservation” In Proceedings of ACM Conference on Management of Data (SIGMOD’06), 2006, pp. 229–240

[39] D. P. M. Kumar and Y. P. Gowramma. “Development of Sensitivity Classification Approach for Personalized Privacy Preservation in Data Publishing (PPPDP)” in International Journal of Innovative Research in Computer and Communication Engineering, 2017

[40] S. Chakraborty, C. Shen, K. R. Raghavan, Y. Shoukry, M. Miller, M. B. Srivastava, “ipShield: a framework for enforcing context-aware privacy” in Proceedings of USENIX Symposium on Networked Systems: Design and Implementation (NSDI) ; 2014; Seattle, WA, pp. 143–156 .

[41] Big Data and the Challenge of Unstructured Data [Online]. Available: https://www.ciklum.com/blog/big-data-and-the-challenge-of-unstructured-data/

[42] V. Vincze, R. Farkas, “De-identification in Natural Language Processing”, in proceedings of 37th Convention of Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014, pp. 1300-1303.

[43] L. Sweeny, “Replacing Personally Identifiable Information in Medical Records, the Scrub System”, in the Journal of the American Medical Informatics Association, 1996.

[44] I. Neamatullah, M. Douglass, L. Lehman, A. Reisner, M. Villarroel, W. Long, G. Clifford, “Automated de-identification of free-text medical records” in BMC Medical Informatics and Decision Making, 2014.

[45] R. Motwani, S. Nabar, “Anonymizing unstructured data” [ONLINE]. Available: https://arxiv.org/pdf/0810.5582.pdf

[46] UK Data Service, [ONLINE]. Available: https://bitbucket.org/ukda/ukds.tools.textanonhelper/wiki/Home

[47] H. Vico, D. Calegari, “Software Architecture for Document Anonymization” in Electronic Notes in Theoretical Computer Science, 2015, pp.83-100.

[48] B. Kleinberg, M. Mozes, Web-based Text Anonymization with Node.js: Introducing NETANOS (Named entity-based Text Anonymization for Open Science) in the Journal of Open Source Software

[49] X. Carreras, L. Marquez, L. Padro´, “A Simple Named Entity Extractor using AdaBoost”, in Proceedings of Proceedings of CoNLL, 2003

[50] Industrial-Strength Natural Language Processing [ONLINE]. Available: https://spacy.io/

[51] Overview of Classification Methods in Python with Scikit-Learn [ONLINE]. https://stackabuse.com/overview-of-classification-methods-in-python-with-scikit-learn/

[53] K. LeFevre, D. DeWitt, R. Ramakrishnan, “Mondrian Multidimensional K-Anonymity”, in Proceedings of 22nd International Conference on Data Engineering (ICDE), 2006

[54] 5 Reasons why you should use Cross-Validation in your Data Science Projects [ONLINE]. Available: https://towardsdatascience.com/5-reasons-why-you-should-use-cross-validation-in-your-data-science-project-8163311a1e79

[55] Understanding Confusion Matrix [ONLINE]. Available: https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

[56] Practice Problem: Twitter Sentiment Analysis [ONLINE]. Available: https://datahack.analyticsvidhya.com/contest/practice-problem-twitter-sentiment-analysis/#ProblemStatement

[57] Twitter Sentiment Analysis [Online]. Available: https://www.kaggle.com/paoloripamonti/twittersentiment-analysis