REFERENCES

[1] A. Selcuk Uluagac & Mauro Conti Abbas Acar, Hidayet Aksu. 2018. A Survey on Homomorphic Encryption Schemes: Theory and Implementation. (2018). https://dl.acm.org/doi/pdf/10.1145/3214303

[2] Martin Albrecht, Melissa Chase, Hao Chen, Jintai Ding, Shafi Goldwasser, Sergey Gorbunov, Shai Halevi, Jeffrey Hoffstein, Kim Laine, Kristin Lauter, Satya Lokam, Daniele Micciancio, Dustin Moody, Travis Morrison, Amit Sahai, and Vinod Vaikuntanathan. 2019. Homomorphic Encryption Standard. Cryptology ePrint Archive, Report 2019/939. https://eprint.iacr.org/2019/939.

[3] Satanjeev Banerjee and Ted Pedersen. 2003. The design, implementation, and use of the ngram statistics package. In *International Conference on Intelligent Text Processing and Computational Linguistics*. Springer, 370–381.

[4] D. Bogdanov. 2008. Sharemind: A framework for fast privacy-preserving computations. *ES-ORJCS 2008* (2008), 192–206. https://ci.nii.ac.jp/naid/20001634171/en/

[5] Luca Bonomi, Li Xiong, Rui Chen, and Benjamin CM Fung. 2012. Frequent grams based embedding for privacy preserving record linkage. In *Proceedings of the 21st ACM international conference on Information and knowledge management*. 1597–1601.

[6] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. 2011. Fully Homomorphic Encryption without Bootstrapping. Cryptology ePrint Archive, Report 2011/277. https://eprint.iacr.org/2011/277.

[7] Hao Chen, Kim Laine, and Peter Rindal. 2017. Fast Private Set Intersection from Homomorphic Encryption. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, Texas, USA) *(CCS '17)*. Association for Computing Machinery, New York, NY, USA, 1243–1255. https://doi.org/10.1145/3133956.3134061

[8] Peter Christen. 2005. Probabilistic data generation for deduplication and data linkage. In *International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 109–116.

[9] Scott D Constable and Steve Chapin. 2018. libOblivious: A c++ library for oblivious data structures and algorithms. (2018).

[10] Nigel P. Smart Craig Gentry, Shai Halevi. 2012. Homomorphic Evaluation of the AES Circuit. (2012). https://link.springer.com/chapter/10.1007%2F978-3-642-32009-5_49

[11] Ronald Cramer, Ivan Damgård, and Jesper B. Nielsen. 2001. Multiparty Computation from Threshold Homomorphic Encryption. In *Advances in Cryptology — EUROCRYPT 2001*, Birgit Pfitzmann (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 280–300.

[12] Cybernetica. 2019. *SecreC (version 2019.03)*. https://sharemind-sdk.github.io/stdlib/reference

[13] Cybernetica. 2019. *Sharemind SDK and Academic server (version 2019.03)*. https://sharemind.cyber.ee/sharemind-mpc

[14] Yvo Desmedt. 2011. *Threshold Cryptography*. Springer US, Boston, MA, 1288–1293. https://doi.org/10.1007/978-1-4419-5906-5_330

[15] L Dusserre, C Quantin, and H Bouzelat. 1995. A one way public key cryptosystem for the linkage of nominal files in epidemiological studies. *Medinfo. MEDINFO* 8 (1995), 644–647.

[16] Fatih Emekci, Divyakant Agrawal, Amr E Abbadi, and Aziz Gulbeden. 2006. Privacy preserving query processing using third parties. In *22nd International Conference on Data Engineering (ICDE'06)*. IEEE, 27–27.

[17] Craig Gentry. 2009. A FULLY HOMOMORPHIC ENCRYPTION SCHEME. (2009). https://crypto.stanford.edu/craig/craig-thesis.pdf

[18] Shai Halevi. 2021. *HElib (version 2.1.0)*. https://github.com/homenc/HElib

[19] Paul Jaccard. 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vaudoise Sci Nat* 37 (1901), 547–579.

[20] Yongsoo Song Jung Hee Cheon & Kyoohyung Han, Andrey Kim & Miran Kim. 2019. A Full RNS Variant of Approximate Homomorphic Encryption. (2019). https://link.springer.com/chapter/10.1007%2F978-3-030-10970-7_16

[21] Yongsoo Song Jung Hee Cheon, Andrey Kim & Miran Kim. 2017. Homomorphic Encryption for Arithmetic of Approximate Numbers. (2017). https://link.springer.com/chapter/10.1007%2F978-3-319-70694-8_15

[22] Jonathan Katz and Yehuda Lindell. 2014. *Introduction to Modern Cryptography, Second Edition*. CRC Press. https://www.crcpress.com/Introduction-to-Modern-Cryptography-Second-Edition/Katz-Lindell/p/book/9781466570269

[23] Gregory W Lesher, Bryan J Moulton, D Jeffery Higginbotham, et al. 1999. Effects of ngram order and training text size on word prediction. In *Proceedings of the RESNA'99 Annual Conference*. Citeseer, 52–54.

[24] Jure Leskovec, Anand Rajaraman, and Jeffrey David Ullman. 2014. *Mining of Massive Datasets* (2nd ed.). Cambridge University Press, USA.

[25] Yehida Lindell. 2005. Secure multiparty computation for privacy preserving data mining. In *Encyclopedia of Data Warehousing and Mining*. IGI global, 1005–1009.

[26] Matthew Michelson and Craig A. Knoblock. 2006. Learning Blocking Schemes for Record Linkage. In *Proceedings of the 21st National Conference on Artificial Intelligence - Volume 1* (Boston, Massachusetts) *(AAAI'06)*. AAAI Press, 440–445. http://dl.acm.org/citation.cfm?id=1597538.1597609

[27] Frank Niedermeyer, Simone Steinmetzer, Martin Kroll, and Rainer Schnell. 2014. Cryptanalysis of basic bloom filters used for privacy preserving record linkage. *German Record Linkage Center, Working Paper Series, No. WP-GRLC-2014-04* (2014).

[28] Olga Ohrimenko, Felix Schuster, Cédric Fournet, Aastha Mehta, Sebastian Nowozin, Kapil Vaswani, and Manuel Costa. 2016. Oblivious multi-party machine learning on trusted processors. In *25th {USENIX} Security Symposium ({USENIX} Security 16)*. 619–636.

[29] Catherine Quantin, Hocine Bouzelat, FAA Allaert, Anne-Marie Benhamiche, Jean Faivre, and Liliane Dusserre. 1998. How to ensure data security of an epidemiological follow-up: quality assessment of an anonymous record linkage procedure. *International journal of medical informatics* 49, 1 (1998), 117–122.

[30] Jaak Randmets et al. 2017. *Programming Languages for Secure Multi-party Computation Application Development*. Ph.D. Dissertation.

[31] Monica Scannapieco, Ilya Figotin, Elisa Bertino, and Ahmed K Elmagarmid. 2007. Privacy preserving schema and data matching. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*. 653–664.

[32] Rainer Schnell, Tobias Bachteler, and Jörg Reiher. 2009. Privacy-preserving record linkage using Bloom filters. *BMC medical informatics and decision making* 9, 1 (2009), 1–11.

[33] SEAL 2020. Microsoft SEAL (release 3.6). https://github.com/Microsoft/SEAL. Microsoft Research, Redmond, WA.

[34] Ziad Sehili, Lars Kolb, Christian Borgs, Rainer Schnell, and Erhard Rahm. 2015. Privacy preserving record linkage with PPJoin. *Datenbanksysteme für Business, Technologie und Web (BTW 2015)* (2015).

[35] Jaydip Sen. 2013. Homomorphic Encryption: Theory & Application. (2013). https://arxiv.org/pdf/1305.5886.pdf

[36] Taffee T Tanimoto. 1958. Elementary mathematical theory of classification and prediction. (1958).

[37] PALISADE team. 2020. *PALISADE Lattice Cryptography Library (release 1.10.6)*. https://palisade-crypto.org

[38] Dinusha Vatsalan, Peter Christen, and Vassilios S Verykios. 2013. A taxonomy of privacy-preserving record linkage techniques. *Information Systems* 38, 6 (2013), 946–969.

[39] Dinusha Vatsalan, Ziad Sehili, Peter Christen, and Erhard Rahm. 2017. Privacy-preserving record linkage for big data: Current approaches and research challenges. In *Handbook of Big Data Technologies*. Springer, 851–895.

[40] Bing Wang, Wei Song, Wenjing Lou, and Y Thomas Hou. 2015. Inverted index based multi-keyword public-key searchable encryption with strong privacy guarantee. In *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2092–2100.

[41] Xiao Shaun Wang, Yan Huang, Yongan Zhao, Haixu Tang, XiaoFeng Wang, and Diyue Bu. 2015. Efficient genome-wide, privacy-preserving similar patient query based on private edit distance. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. 492–503.

[42] Andrew C Yao. 1982. Protocols for secure computations. In *23rd annual symposium on foundations of computer science (sfcs 1982)*. IEEE, 160–164.

[43] Samee Zahur and David Evans. 2015. Obliv-C: A Language for Extensible Data-Oblivious Computation. *IACR Cryptol. ePrint Arch.* 2015 (2015), 1153.

[44] Eric Zhu. 2017. *Datasketch (version 1.2.5)*. https://github.com/ekzhu/datasketch

[45] Ruiyu Zhu and Yan Huang. 2017. Efficient privacy-preserving general edit distance and beyond. *IACR Cryptology ePrint Archive* 2017 (2017), 683.