

# CSCM21: Designing in Trust, Understanding, and Negotiation - Coursework: BeatLonliNess plc

Andy Gray  
445348

28/04/2021

## **1 Based on the scenario summarised above and the aspects of responsible design learnt in the lectures, discuss issues (legal, ethical, and technological) with the business proposal of BeatLonliNess plc. (13 Marks)**

In legal aspects, the BeatLonliNess (BLN) plc platform must abide by the data protection and GDPR rules that the British government and the EU have set out. These are rules to ensure that companies keep their users' data safe and secure while also holding the user's information relevant to the organisation. BLN could achieve this from the word go as they start to expand by using privacy by design method. Privacy by design ensures that privacy is a requirement in the design and development process that includes encryption, federated learning, differential privacy, access control, transparency, and finally, consent [?]. The Information Commissioner's Office (ICO) states that policies and procedures are needed to get implemented to ensure data protection issues get considered when systems, services, products and business practices involving personal data are designed and implemented. Therefore, as a result, personal data gets protected by default, ensuring that safeguarding individuals' rights. These rights include data minimisation, pseudonymisation and purpose limitation [?].

With BLN using AI-supported algorithms, they must get designed to be reliable. For the algorithms to be reliable, they would also need to carry out their tasks with high accuracy, ensuring that the generated results are as expected to get generated. Therefore it would be a good idea for BLN to build their algorithms with explainability within them. Making the algorithm explainable would allow the users to trust the algorithms more and see what factors impact their matchings. However, the data must not create any potential bias within the models to allow the matches to happen effectively, but removing bias is challenging to spot and remove. BLN must remove any potential bias from their datasets to ensure that no member of the platform gets discriminated against, whether it be because of their gender, race, religion, ethnicity or skin colour.

These issues also lead to BLN making sure that the algorithms and models they use are also responsible AI. For the AI to be responsible, the models will need to be transparent, trustworthy, ethical and respecting the privacy of the users.

- 2 Summarise relevant examples of related media coverage in the last 5 years. (7 Marks)
- 3 Use the ETHICS GUIDELINES FOR TRUSTWORTHY AI published by AI high-level expert group of the European Commission in April 2019, in particular the TRUSTWORTHY AI ASSESSMENT LIST (p.24 of the report and standalone document), to discuss requirements for the system proposed by Beat-LonliNess plc. (Links to these documents are posted with the assessment brief.) (10 Marks)

## References

- [1] ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z., CITRO, C., CORRADO, G. S., DAVIS, A., DEAN, J., DEVIN, M., GHEMAWAT, S., GOODFELLOW, I., HARP, A., IRVING, G., ISARD, M., JIA, Y., JOZEFOWICZ, R., KAISER, L., KUDLUR, M., LEVENBERG, J., MANÉ, D., MONGA, R., MOORE, S., MURRAY, D., OLAH, C., SCHUSTER, M., SHLENS, J., STEINER, B., SUTSKEVER, I., TALWAR, K., TUCKER, P., VANHOUCKE, V., VASUDEVAN, V., VIÉGAS, F., VINYALS, O., WARDEN, P., WATTENBERG, M., WICKE, M., YU, Y., AND ZHENG, X. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] CHO, K., VAN MERRIËNBOER, B., GULCEHRE, C., BAHDANAU, D., BOUGARES, F., SCHWENK, H., AND BENGIO, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078* (2014).
- [3] CHUNG, J., GULCEHRE, C., CHO, K., AND BENGIO, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555* (2014).
- [4] GEEKS FOR GEEKS. Understanding of lstm networks, 2020.
- [5] GÉRON, A. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, 2019.
- [6] GREFF, K., SRIVASTAVA, R. K., KOUTNÍK, J., STEUNEBRINK, B. R., AND SCHMIDHUBER, J. Lstm: A search space odyssey. *IEEE transactions on neural networks and learning systems* 28, 10 (2016), 2222–2232.
- [7] HOCHREITER, S., AND SCHMIDHUBER, J. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [8] KAGGLE. Real or not? nlp with disaster tweets, 2020.
- [9] LIN, Y. 10 twitter statistics every marketer should know in 2021, 2020.
- [10] LOPER, E., AND BIRD, S. Nltk: The natural language toolkit. In *In Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics. Philadelphia: Association for Computational Linguistics* (2002).
- [11] MCKINNEY, W. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* (2010), S. van der Walt and J. Millman, Eds., pp. 51 – 56.
- [12] MONAGHAN, M. Website load time statistics: Why speed matters in 2020, 2020.
- [13] OLAH, C. Understanding lstm networks, 2015.

- [14] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M., AND DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
- [15] VISUAL OAK. Internet facts, history, and resources for 2020, 2020.