## GLOSSARY/TERMINOLOGY

1) **Counterfactual fairness –** AI decision is fair towards an individual if it is the same in (a) the actual world and (b) a counterfactual world where the individual belonged to a different demographic group [3]
2) **CI/CD [2]** – collection of software engineering practices enabling delivery of software in Agile manner characterised by high-frequency deployments of changes to systems and strong focus on automating aspects of delivery. The key phases include:
    a. **Continuous integration** - An approach to integrating, building, and testing code within the software development environment.
    b. **Continuous delivery** - An approach to software development in which software can be released to production at any time. Frequent deployments are possible, but deployment decisions are taken case by case, usually because organizations prefer a slower rate of deployment.
    c. **Continuous deployment** - An approach to software development in which changes go through the pipeline and are automatically put into the production environment, enabling multiple production deployments per day. **Continuous deployment** relies on **continuous delivery**.
3) **COBIT (Control Objectives for Information Technologies)** – high-level IT governance and control framework established by ISACA professional organization[1]. Has been developed as an open standard and is now increasingly adopted globally as the control and management framework for effective IT governance [14]
4) **CRISP-DM (Cross-Industry Standard Process for Data Mining)** – structured process/methodology to execute data mining projects [4], released in 1999 and now the most widely used data mining model
5) **GDPR (General Data Protection Regulation)** – a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA), in force from 25 May 2018
6) **Data biases [5]–** inherent biases in the datasets which may lead to discriminatory or other undesired modelling outcomes**.** There are two common types of **data biases**:
    a. **Selection bias -** occurs when the data used to produce the model are not fully representative of the actual data or environment that the model may receive or function in, e.g. **omission bias** (omission of certain characteristics from the dataset) or **stereotype bias** (attributing certain characteristics from the dataset to wider population)
    b. **Measurement bias -** data is systematically skewed in a particular direction, for example, could be caused by data collection method or device.
7) **[Data] 'processing'** - means any operation or set of operations which is performed on personal data or on sets of personal data, whether by automated means (GDPR, Art. 4)
8) **De-biasing** – reduction, mitigation of bias
9) **Fairness (in data mining and machine learning context) [10]** – can be defined in terms of the two conditions that have to be met concurrently:
    a. Condition 1 (relates to direct discrimination) - people that are similar in terms of non-protected characteristics should receive similar predictions, and

[1] Information Systems Audit and Control Association

b. Condition 2 (relates to indirect discrimination) - differences in predictions across groups of people can only be as large as justified by their non-protected characteristics[2]

*Example to satisfy Condition 1 (twin test illustration from [10]):* Gender is the protected attribute, and there are two identical twins who share all the characteristics except gender. Test is passed if both receive identical predictions.

*Example to satisfy Condition 2 (so-called 'red-lining' practice from [10]):* Banks denied loans for residents of selected neighbourhoods (non-white dominated) more compared to other neighbourhoods. Thus, even though race was not used as a decision criterion and people of different races ("twins") from the same neighbourhood were treated equally, groups of similar people (non-white) were treated differently. In particular, non-white population on overall level was negatively affected by lower positive decision rates in the non-white-dominated neighbourhoods

10) **ITIL (Information Technology Infrastructure Library)** – best practices framework for IT service management [15]. Initially published by the UK government, now have become the most widely adopted IT Service Management (ITSM) framework and is regarded as de-facto standard in private and public sectors around the world [14]
11) **Model explainability** - achieved by explaining how deployed AI models' algorithms function and/or how the decision-making process incorporates model predictions [5]
12) **Model repeatability** – model's ability to consistently perform an action or make a decision, given the same scenario [5]
13) **Output (as per CRISP-DM)** – The tangible result of performing a task [1]
14) **(Pseudo)anonymization** - means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information (GDPR, Art. 4). In case of pseudoanonymization (re)identification is still possible, while in case of full anonymization data cannot be re-identified
15) **Phase (as per CRISP-DM)** – A term for the high-level part of the CRISP-DM process model; consists of related tasks [1]
16) **Reference model (as per CRISP-DM)** – Decomposition of data mining projects into phases, tasks, and outputs [1]
17) **Requirements [7]**- a specification of what should be implemented, in particular, descriptions of how he system should behave, or of a system property or attribute
18) **Requirements Engineering [7]**– the process of discovering, documenting, and managing the requirements for a computer-based system. The goal of RE is to produce a set of system requirements which, as far as possible, is complete, consistent, relevant and reflects what the customer actually wants
19) **Requirement Engineering activities [5]-[8]** include:
    a. **Requirements development** subdivided into elicitation (needs discovery), analysis, specification (documenting), and validation
    *NB! These activities are underpinned with stakeholders' negotiation*

[2] Both these conditions are commonly used to define fairness and are known as Lipschitz condition and statistical parity respectively [10] (discussed in [10] and originally presented in [11] in the context of classification). These two measures relate to assessing fairness of the decisions. There also other measures that focus to assess fairness of the assigned scores/predictions, see for example [12]. Also, multiple notions of fairness sometimes are not possible to satisfy (for more guidance please see for example [13])

b. **Requirements management** involves controlling requirements changes (tracing, documenting, validating, etc.)

20) **Robustness** - ability of a computer (AI-based) system to cope with errors during execution and erroneous input [5]
21) **Software Engineering (Development) life-cycle** - framework defining tasks performed at each step in the software development process and followed by a development team within the software organization.
22) **Task (as per CRISP-DM)** – A series of activities to produce one or more outputs; part of a phase [1]

**REFERENCES:**


*[1] Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, Wirth R. “Crisp-dm 1.0 step-by-step data mining guide”. Chicago: SPSS Inc., 2000.*

*[2] ITIL®4: Create, Deliver and Support, Axelos, TSO, London, 2020.*

*[3] Kusner, M. J., Loftus, J., Russell, C., and Silva, R., “Counterfactual fairness”. In Advances in neural information processing systems, pp. 4066-4076, 2017.*

*[4] Marban, O., Mariscal, G., Segovia, J. “A data mining and knowledge discovery process model. Data Mining and Knowledge Discovery in Real Life Applications”, edited by P. Julio and K. Adem, pp. 438-453, Paris, I-Tech, Vienna, Austria, 2009.*

*[5] Singapore’s Info-communications Media Development Authority (“IMDA”) and Personal Data Protection Commission (“PDPC”), “Model AI Governance Framework”, 2nd Edition, Jan 2020.*

*[6] Sommerville, Ian. "Integrated requirements engineering: A tutorial." IEEE software 22.1 (2005): 16-23.*

*[7] Sommerville, I., and Sawyer, P. “Requirements engineering: a good practice guide.” John Wiley & Sons, Inc., 1997.*

*[8] Wiegers, Karl. ”More about software requirements: thorny issues and practical advice”. Microsoft Press, 2005.*

*[9] Bourque, P., and Fairley, R. E. Guide to the software engineering body of knowledge (SWEBOK (R)): Version 3.0. IEEE Computer Society Press, 2014.*

*[10] Žliobaitė, I., “Measuring discrimination in algorithmic decision making”, Data Mining and Knowledge Discovery, 31(4), pp.1060-1089, 2017.*

*[11] Dwork, C., Hardt, M., Pitassi, T., Reingold, O., Zemel, R.S.”Fairness through awareness”. In: Proc. of Innovations in Theoretical Computer Science, pp. 214-226, 2012.*

*[12] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, Aziz Huq, “Algorithmic Decision Making and the Cost of Fairness”. KDD 2017: 797-806*

*[13] Pessach, Dana, and Erez Shmueli. "Algorithmic fairness." arXiv preprint arXiv:2001.09784 (2020).*

*[14] Guldentops, Erik, Tony Betts, and Gary Hodgkiss. "Aligning COBIT, ITIL and ISO 17799 for Business Benefit." Office of Government Commerce, Norwich, UK2000 (2007).*

*[15] Arraj, Valerie. "ITIL®: the basics." Buckinghampshire, UK (2010).*