Fairness Metrics and Maximum Completeness for the Prediction of Discrimination

Alessandro Simonetta[1], Tsuyoshi Nakajima[2], Maria Cristina Paoletti[3] and Alessio Venticinque[4]

[1]*Department of Enterprise Engineering University of Rome Tor Vergata, Rome, Italy*
[2]*Department of Computer Science and Engineering Shibaura Institute of Technology, Tokyo, Japan*
[3]*Rome, Italy*
[4]*Naples, Italy*

Abstract
Data has assumed increasing importance within the global economy, and its use is becoming more pervasive in multiple contexts. However, learning systems are exposed to various critical issues that can be addressed through ISO standards. Indeed, machine learning (ML) models may be exposed to the risk of perpetrating societal prejudice simply because the same bias exists in the data. Based on these notions, we have build a model to identify similar treatment groups based on the type of classification errors made by ML algorithms. A way to calculate fairness indices on the protected attributes of the dataset will be illustrated in the article. Finally, we will consider the degree of relationship existing between maximal completeness and fairness of forecasting algorithms through an inverse procedure of constructing a complete dataset. The use of mutual information provided an alternative method for calculating synthetic fairness indices and a useful basis for future research.

Keywords
fairness, machine learning, maximum completeness, treatment similarity, mutual information, entropy

1. Introduction

Data has become increasingly important within the global economy, and its use, which often occurs through sophisticated learning systems, is becoming more pervasive in many areas.
The Economist [1] was one of the first to define data the oil of the modern age. With the rise of Artificial Intelligence (AI) algorithms in decision support, data quality has become always more important, therefore Forbs [2] points to data as the fuel of ML algorithms. Consequently, a new business has emerged based on their collection and sale. WEB giants such as Google offer free services and products with the target of collecting information often for advertising purposes. Many social platforms are free, and companies earn considerable sums from selling the information rather than from payment services. This has pushed these companies to use increasingly sophisticated technologies [3] and algorithms to collect information and integrate it with those from other data sources to maximize their insights. In addition, as presented in the documentary *"The Social Dilemma"* [4] information about users, including contacts and interactions on platforms are being used to influence their behavior without them realizing it, by providing the right ad hoc inputs.

While being able to have multiple data makes it possible to perform analyses on phenomena, we must consider that ML algorithms, as in [5] and [6], are affected by the completeness and redundancy of information to train them. The presence of bias within them can cause discrimination regarding ethnic, gender, religious, race, and cultural minorities, etc. An emblematic example is the one related to the *Compas* dataset [7] where the algorithm used to predict inmates recidivism unfairly disfavored people of *African-American race*.

In the last few years, attention to data quality and its use has increased, and, especially in Europe, legislators have become aware of the existence of the problem [8]. It is worth mentioning that also in the *General Data Protection Regulation* (GDPR) 2016/679 [9], defined to harmonize the data privacy laws among the European countries, there are data quality notions such as accuracy, timeliness and security. The same could be found in the European regulation *Solvency II* [10], which states the need for insurance companies to have internal procedures and processes in place to ensure the appropriateness.

As we mentioned in [11] we believe that a good solution to ensure the correct use of data and their quality according to regulation and ethics values is the compliance to ISO standards: ISO/IEC 27000 [12], ISO 31000 [13] e ISO/IEC 25000 [14]. The introduction of maximum completeness, as dataset balance index, and its relation to fairness metrics are emanations of the of SQuaRE approach in measuring data quality and assessing its implications.

Woodstock'21: Symposium on the irreproducible science, June 07–11, 2021, Woodstock, NY
alessandro.simonetta@gmail.com (A. Simonetta); tsnaka@shibaura-it.ac.jp (T. Nakajima); mariacristina.paoletti@gmail.com (M. C. Paoletti)
0000-0003-2002-9815 (A. Simonetta); 0000-0002-9721-4763 (T. Nakajima); 0000-0001-6850-1184 (M. C. Paoletti); 0000-0003-3286-3137 (A. Venticinque)

2. The Present Situation

Although it is difficult to estimate the cost of the absence of quality in data, a primary goal for organizations (public and private) that base their business on the digitization of processes and the operation of the organization itself is to have trusted data [15]. Some experiences show how the application of the *SQuaRE series* is a solution for measuring and monitoring data quality over time. In Italy, the first indication towards public administration managing databases of national interest was in 2013, in fact the *Agency for Digital Italy* (AgID) had identified in the ISO /IEC 25012 standard the data quality model to be adopted [16].

Since in 2013, AgID had identified within the 15 quality characteristics, those that should be inescapably used (accuracy, consistency, completeness, and newness) for databases of national interest. In the three-year plan for public administration information technology 2021-2023 [17], AgID confirms increasing data and metadata quality as a strategic goal (OB2.2).

In [18] are reported three case studies of data quality evaluation and certification process about repositories. The different visions are analyzed to evaluate the impact of the adoption of the ISO/IEC 25012, ISO/IEC 25024 and ISO/IEC 25040 and their benefit recognized in the three organization before and after the process. The results show that applying their methodology helps the organization to get a better sustainability in the long term, improve the knowledge of the business and drive the organizations in better data quality initiatives for the future.

Among the environments in which the above ISO standards can be most useful are undoubtedly those where the information contains sensitive or safety data [19] such as the healthcare and legal domains. An example is the proposed OpenEHR standard in [20]. The issues that touch clinical records from the perspective of data quality are presented in [21]. In [22] the authors propose a generalized model for big data: a solution based on the application of ISO/IEC 25012 and ISO/IEC 25024. The study introduces three data quality dimensions: Contextual Consistency, Operational Consistency and Temporal Consistency. In [11] the authors show how using the SQuaRE series can ensure GDPR compliance. In [23] the study examines discrimination against nonwhite teachers who are present on online English language teaching platforms.

One possible solution to the problem that bias in the data can propagate into the inferences of ML algorithms is through the dataset labeling mechanism presented in [24]. In [25] the authors present a range of fair access and information presentation (quantity of data presented to the user and order of priority) issues that may affect the fairness of computing systems. Although these issues are related to the biases within the data, characteristics of recommender systems can introduce a greater degree of uncertainty. These are related to the permissions of the users who use them to access the information or the size of the data that can be processed by the algorithms. This makes it even more difficult to find countermeasures to avoid discrimination.

Finally, in [26] the authors show a methodology for identifying critical attributes that can lead to discrimination by classification-based learning systems.

3. Solution Proposed

When using an ML-based recommendation system on a dataset where bias is present, the bias propagates within the model itself, replicating the guesswork and prejudices in the data. So, we run the risk of thinking that we applied an objective and neutral evaluation system, while we are using a biased system within an AI algorithm.

One of the purposes of this research is to verify that the system behaves in a non-discriminatory way toward certain groups. By considering the different fairness measures in [27], it is possible to calculate their value with respect to two groups, identified by a protected attribute, to see if there are any disparities in treatment. For example the formal criterion of *Independence* requires that the sensitive attribute A would be statistically independent of the predicted value R and this could be calculated $\forall i, j : i \neq j$ as:

$$P(R = 1|A = a_i) = P(R = 1|A = a_j) \tag{1}$$

To understand whether an attribute is a cause of discrimination in prediction outcomes, that is, whether there are homologous treatment groups, it is necessary to know the attribute's level of fairness. Ideally, therefore, should be better to have a single measure that gives an idea of how likely that attribute is to lead to discrimination.

In [26], the authors propose a method to compute several synthetic indices related to the fairness of the classification system. Two different methods are described in the article: the first performs clustering with *DBSCAN* and *Kmeans* methods while the second, *MaxMin*, searches for the worst case by dividing the protected attribute instances into privileged and unprivileged. Both methods allow grouping the elements of a protected attribute according to the type of treatment. In this way the calculation of the synthetic index is based on a few influence classes returning to the definition from which we started [27]. These two approaches were used to test for a link between the notion of *maximum completeness* and *fairness indices*. This would allow a priori identification of

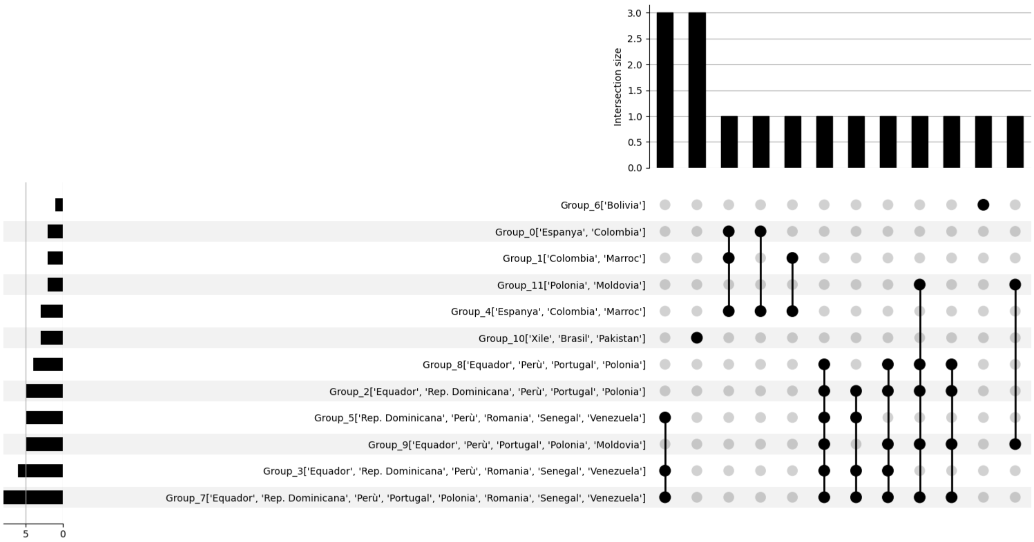

Figure 1: Metropolitan Diagram related to nationalities in Juvenile Dataset with groups intersections

whether learning on a present dataset can lead to minority discrimination. At this point, alternative methods are proposed to identify homogeneous treatment groups with respect to the result obtained from a classification system. Algorithms may err toward some groups equally, i.e. for African-Americans and Native Americans they may give a degree of recidivism in excess of what happens in reality.

3.1. Identification of homogeneous treatment groups

To start, we need to calculate the fairness indices reported in [27] for the protected attributes of the dataset, considering the predictions of the classification algorithm and the actual corrected result.

We refereed to a classic case study for this kind of problem: the Compas dataset [7], where we observed a similar trend between groups. Table 1 shows the values of the 6 fairness indices for the protected attribute *Race*: Independence (Ind), Separation True Positive Rate (SepTPR), Separation False Positive Rate (SepFPR), Sufficiency Positive Predictive Value (SufPPV), Sufficiency Negative Predictive Value (SufNPV) and Overall Accuracy Equality (OAE).

Table 2 shows the correlation matrix according to Pearson's coefficient and the existence of correlation between the indices measured for different ethnicities. Considering a correlation value of 0.9, it is easy to detect the existence of two treatment groups (Table 3): G0 and G1.

Although this method works well for the case study, the issue becomes more complicated when there are categorical attributes with higher cardinality as the number of relations increases. With reference to the *Juvenile* dataset [28], considering the *V3_nacionalitat* attribute representing the nationality of the students, it is possible to draw the phenomenon through a subway diagram (Fig. 1). In this graph, it is easier to check intersections between sets. For example, Group 0 and Group 1 have the element Colombia in common. The top histogram shows the number of elements participating in the intersection while the left histogram shows the number of elements in the group.

The result obtained with the Pearson coefficient threshold of 0.9 identifies twelve homogeneous treatment groups. In order to reduce their number, we kept as a representative of a set of groups the one that contained them in the inclusion relation. This reduced the twelve to four completely disjointed groups.

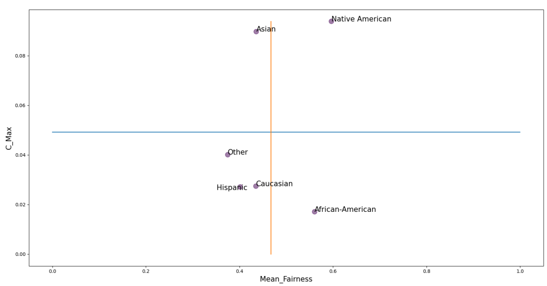

Figure 2: Scatterplot of races in Compas Dataset, mean of fairness metrics Vs maximum completeness

Table 1
Table of fairness measures

Race	Fairness Index					
	Ind.	*SepTPR*	*SepFPR*	*SufPPV*	*SufNPV*	*OAE*
Caucasian	33,10%	50,36%	22,01%	59,48%	29,00%	67,19%
Hispanic	27,70%	41,80%	19,38%	56,03%	29,89%	66,21%
Other	20,41%	33,87%	12,79%	60,00%	30,04%	67,93%
Asian	22,58%	62,50%	8,70%	71,43%	12,50%	83,87%
African American.	57,61%	71,52%	42,34%	64,95%	35,14%	64,91%
Native American.	72,73%	100%	50%	62,50%	0,00%	72,73%

Table 2
Table of Pearson Cefficient Correlation

Race	Race					
	African-A.	*Native A.*	*Caucasian*	*Hispanic*	*Other*	*Asian*
African-American	1	0,901	0,801	0,680	0,562	0,848
Native American	0,901	1	0,515	0,364	0,210	0,596
Caucasian	0,801	0,515	1	0,983	0,941	0,991
Hispanic	0,680	0,364	0,983	1	0,984	0,956
Other	0,562	0,210	0,941	0,984	1	0,900
Hispanic	0,848	0,596	0,991	0,956	0,900	1

Table 3
Groups Aggregation

Groups	Races			
	1	*2*	*3*	*4*
G0	American African	Native American		
G1	Caucasian	Hispanic	Other	Asian

3.2. Relationship between mean of fairness indexes and C_{MAX}

At this point, we studied if there was a relationship between the composition of the groups made using Pearson's coefficient and the maximum completeness, as shown in the [29], [30] studies. For this purpose, we used the scatterplot diagram in which each ethnicity was drawn in relation to the pair of values: mean of the fairness indices, in the abscissa, and maximum completeness, in the ordinate (Fig. 2). Considering the positioning of the different ethnic groups and a scale that reports the highest value as the limit of the diagram, we observe that they tend to cluster on average around the grand mean of the fairness attributes, most noticeably when we look at the privileged group. Items belonging to the same group tend to remain close relative to the fairness index. These considerations are less true for the *maximum completeness index* (C_{MAX}).

After extending the analysis to the different attributes of the datasets already present in [26] [31], C_{MAX} seems to be a strongly characterizing parameter, more so than the other indices proposed in [32]. In fact, repeating the analysis on other protected attributes, such as *V3_nacionalitat* of the Juvenile dataset, within the scatterplot the clustering of similarly treated elements was found to be strongly related not only to the average of the fairness indices, but also to the C_{MAX} as present in Fig. 3, considering the groups with intersection present in 1.

3.3. Alternative synthetic indices

The presence of outliers in the values of fairness indices related to a protected attribute could impact the valuation of these parameters. For this reason, in this paper we propose a different way of calculating fairness indices. In this research, we calculate independence, separation, sufficiency and OAE using the notion of entropy and *mutual information.* The idea is to find a new representation of the synthetic index that would allow more confident identification of whether a given protected attribute could lead to possible discrimination. Considering the condition of Independence between two groups $A = a_i$ and $A = a_j$:

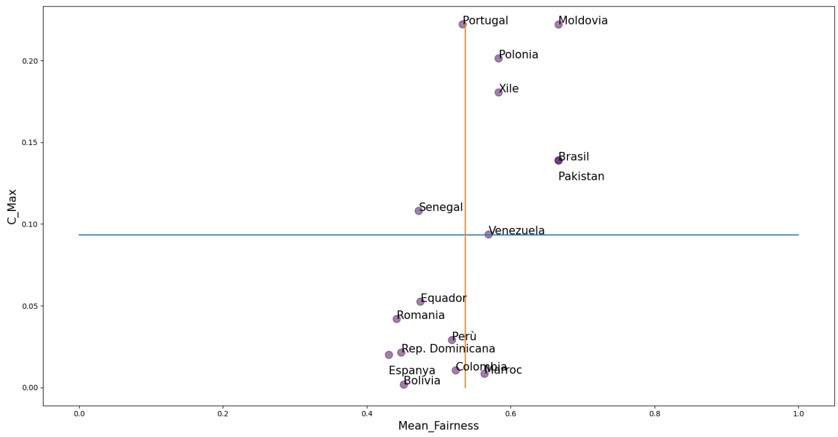

Figure 3: Scatterplot of nationalitis in Juvenile Dataset, mean of fairness metrics Vs maximum completeness

$$|P(R=1|A=a_i)-P(R=1|A\neq a_i)|<\varepsilon \tag{2}$$

This condition can be extended to all categories of the protected attribute and also expressed by orthogonality between the predicted value R and the group A through mutual information. Given two variables, they are independent if their mutual information is zero:

$$I(R,A)=0 \tag{3}$$

remember that mutual information is calculated by the equation:

$$I(R,A)=H(R)+H(A)-H(A,R) \tag{4}$$

where H(R) is the entropy associated with the R function. Thus, the individual terms for calculating independence are:

$$H(R)=\sum_{i=1}^{n}P(r_i)log(P(r_i)) \tag{5}$$

H(A) is the entropy associated to A and it is calculated as:

$$H(A)=\sum_{i=1}^{n}P(a_i)log(P(a_i)) \tag{6}$$

finally, the third term H(A,R) is computed by:

$$H(R,A)=\sum_{i=1,j=1}^{n,m}P(r_i\cap a_j)log(P(r_i\cap a_j)) \tag{7}$$

The other indices can also be expressed by mutual information and in particular referring to [33] and [26] Separation is calculated by:

$$I(R,A|Y)=H(R,Y)+H(A,Y)-H(R,Y,A)-H(Y) \tag{8}$$

sufficiency is expressed by the following equation:

$$I(Y,A|R)=H(Y,R)+H(A,R)-H(Y,R,A)-H(R) \tag{9}$$

finally, the OAE is computed by:

$$\begin{aligned}H(R,A|Y=R)&=H(R,Y=R)+\\&+H(A,Y=R)-H(R=Y,A|R=Y)\end{aligned} \tag{10}$$

Once the mutual information is calculated for the fairness metrics considered, we compared these values with those obtained by applying the methods presented in [26] that refer to calculating MaxMin and clustering with DBSCAN. To achieve this, a normalization process obtained by dividing the different quantities by the maximum achievable value was necessary. In the present

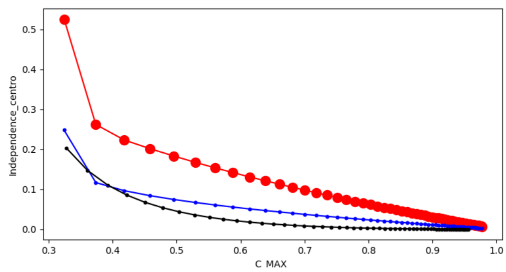

Figure 4: Comparision Mutual Information, DBSCAN and MaxMin methods for Independence

study, we performed the comparison of the three methodologies (MaxMin, DBSCAN and mutual information) applied on the Compas dataset with respect to fairness indices. Since the results show similar trends releted to the fairness measures, without loss of generality, we have reported only the relationships between Independence measure and maximum completeness. In Fig. 4 in red is shown the dependence curve related to MaxMin methodology, in blue that with DBSCAN and in black that with mutual information. The graph, highlighted in Fig. 4 shows the trend of independence versus varying maximum completeness. The process of construction of the dataset initially select few tuples of the original one (C_{MAX}=0.324) and after insert new tuples until the dataset reaches the overall completeness (C_{MAX} = 1), which corresponds to maximum independence.

The curve related to the MaxMin method initially hires greater values than the other two methods, while the phenomenon decreases as the number of records entered increases. Thus, we can conclude from the present research that there is a greater sensitivity of independence measure with respect to varying maximum completeness if the MaxMin method is used.

3.4. Limit and Future Works

This work identified homologous treatment groups using Pearson's coefficient, which detects the correlation between fairness characteristics associated with different groups.

In the future, further research should be done to investigate new similarity mechanisms based on ML and Deep Learning algorithms considering other clustering methodologies that can avoid overlapping between groups.

A second line of research will aim to identify discrimination caused by belonging to more than one protected attribute such as gender and race simultaneously.

Since we do not considered explainable AI algorithms, future works could be extended considering framework that analyze how AI models make decisions (i.e. Watson OpenScale [34]).

4. Conclusion

The use of AI and ML in the decision-making process of many recommendation systems makes it possible to mitigate the risk of subjective classifications.
While these systems are reliable forecasting tools, they do not always allow for an explanation of why such conclusions were reached. Thus, the presence of incomplete or unbalanced data, that can be measured through the SQuaRE series (completeness measures), can lead to biased results.

This work made it possible to us, to calculate similar groups in terms of equivalence of treatment through the application of Pearson's coefficient to synthetic indices related to protected attributes. In such a way, we identified similarities that previously remained hidden in search of possible discrimination.

The other achievement was that we were able to associate fairness measures with protected attributes, independently of those of individual values, using the concept of mutual information and entropy. This approach laid the foundation for new experimentation to relate the response of these measures to changes in maximum completeness.

Finally, we compared the classical approaches [31] versus the method using mutual information and entropy. In this way, we tested the response of fairness measures against maximum completeness and found confirmation against the premises of the work, namely, that non-quality in the data leads to unfair treatments if AI and ML are used in the decision-making process of recommender systems.

References

[1] The Economist, The world's most valuable resource is no longer oil, but data, The Economist, USA (6th May 2017).

[2] B. Marr, The 5 biggest data science trends in 2022, Oct 2021. URL: https://www.forbes.com/sites/bernardmarr/2021/10/04/the-5-biggest-data-science-trends-in-2022/?sh=22f5fc1d40d3.

[3] R. Giuliano, The next generation network in 2030: Applications, services, and enabling technologies, in: 2021 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2021, pp. 294–298. doi:10.23919/EECSI53397.2021.9624241.

[4] J. Orlowski, The social dilemma, Sep. 2020. URL: https://www.netflix.com/it/title/81254224.

[5] G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Re, A. Ricci, S. Spanò, An fpga-based multi-agent reinforcement learning timing synchronizer, Computers and Electrical Engineering 99 (2022) 107749. doi:https://doi.org/10.1016/j.compeleceng.2022.107749.

[6] G. C. Cardarilli, M. Re, L. Di Nunzio, A pseudo-softmax function for hardware-based high speed image classification, Scientific Reports (2021). doi:10.1038/s41598-021-94691-7.

[7] J. Larson, S. Mattu, L. Kirchner, J. Angwin, Compas recidivism dataset, 2016. URL: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

[8] Council of Europe, Recommendation CM/Rec(2020)1 of the Committee of Minis-

ters to member States on the human rights impacts of algorithmic systems (2020).

[9] European Parliament and Council of Europe, Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation) (2016).

[10] Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), 2009.

[11] A. Simonetta, M. C. Paoletti, A. Venticinque, Using the SQuaRE series as a guarantee for GDPR compliance, Ceur-WS 3115 (2021). URL: http://ceur-ws.org/Vol-3114/paper-05.pdf.

[12] International organization for standardization, ”iso/iec 27000:2018, “information technology, security techniques, information security management systems,overview and vocabulary”, 2018. URL: https://www.iso.org/standard/73906.html(accessedNov,2021).

[13] International organization for standardization, ”iso 31000:2018(en) risk management — guidelines”, 2018. URL: https://www.iso.org/iso-31000-risk-management.html(accessedNov,2021).

[14] International organization for standardization, ”iso/iec 25000:2014, “systems and software engineering — systems and software quality requirements and evaluation (square) — guide to square”, 2014.

[15] F. Fallucchi, M. Gerardi, M. Petito, E. De Luca, Blockchain framework in digital government for the certification of authenticity, timestamping and data property, 2021. doi:10.24251/HICSS.2021.282.

[16] ???? URL: https://www.agid.gov.it/sites/default/files/repository_files/circolari/dt_cs_n.68_-_2013dig_-regole_tecniche_basi_dati_critiche_art_2bis_dl_179\-2012_sito.pdf.

[17] Three-year plan for information technology (piano triennale per l’informatica), 2022. URL: https://www.agid.gov.it/it/agenzia/piano-triennale(Access06-22).

[18] F. Gualo, M. Rodriguez, J. Verdugo, I. Caballero, M. Piattini, Data quality certification using iso/iec 25012: Industrial experiences, Journal of Systems and Software 176 (2021) 110938. doi:https://doi.org/10.1016/j.jss.2021.110938.

[19] A. A. Jaber, R. Bicker, The optimum selection of wavelet transform parameters for the purpose of fault detection in an industrial robot, in: 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014), 2014, pp. 304–309. doi:10.1109/ICCSCE.2014.7072735.

[20] M. Sousa, D. Gonçalves-Ferreira, C. Pereira, G. Bacelar, S. Frade, O. Pestana, R. Cruz-Correia, openehr based systems and the general data protection regulation (gdpr), Studies in health technology and informatics 247 (2018) 91–95.

[21] V. C. Pezoulas, K. D. Kourou, F. Kalatzis, T. P. Exarchos, A. Venetsanopoulou, E. Zampeli, S. Gandolfo, F. Skopouli, S. De Vita, A. G. Tzioufas, D. I. Fotiadis, Medical data quality assessment: On the development of an automated framework for medical data curation, Computers in Biology and Medicine 107 (2019) 270–283. URL: https://www.sciencedirect.com/science/article/pii/S0010482519300733. doi:https://doi.org/10.1016/j.compbiomed.2019.03.001.

[22] I. Caballero, M. Serrano, M. Piattini, A data quality in use model for big data, in: M. Indulska, S. Purao (Eds.), Advances in Conceptual Modeling, Springer International Publishing, Cham, 2014, pp. 65–74.

[23] N. M. Curran, Discrimination in the gig economy: the experiences of black online english teachers, Language and Education 0 (2021) 1–15. doi:10.1080/09500782.2021.1981928.

[24] E. Beretta, A. Vetro, B. Lepri, J. C. De Martin, Ethical and Socially-Aware Data Labels: 5th International Conference, SIMBig 2018, Lima, Peru, September 3–5, 2018, Proceedings, 2019, pp. 320–327. doi:10.1007/978-3-030-11680-4_30.

[25] M. D. Ekstrand, A. Das, R. Burke, F. Diaz, Fairness in information access systems, Foundations and Trends® in Information Retrieval 16 (2022) 1–177. doi:10.1561/1500000079.

[26] A. Simonetta, M. C. Paoletti, A. Venticinque, The use of maximum completeness to estimate bias in ai based recommendation systems, SYSTEM 2022 (In Press).

[27] S. Barocas, M. Hardt, A. Narayanan, Fairness and machine learning, 2020. URL: https://fairmlbook.org/, chapter: Classification.

[28] Department of Justice, Recidivism in juvenile justice, 2016. URL: https://cejfe.gencat.cat/en/recerca/opendata/jjuvenil/reincidencia-justicia-menors/index.html.

[29] A. Simonetta, M. C. Paoletti, Designing digital circuits in multi-valued logic, International Journal on Advanced Science, Engineering and Information Technology 8 (2018) 1166–1172. URL: http://ijaseit.insightsociety.org/index.php?option=com_content&view=article&id=9&Itemid=1&article_id=5966. doi:10.18517/ijaseit.8.4.5966.

[30] A. Simonetta, A. Vetrò, M. C. Paoletti, M. Torchiano, Integrating square data quality model with iso 31000 risk management to measure and mitigate

software bias, CEUR Workshop Proceedings (2021) pp. 17–22.
[31] A. Vetrò, M. Torchiano, M. Mecati, A data quality approach to the identification of discrimination risk in automated decision making systems, Government Information Quarterly 38 (2021) 101619. doi:https://doi.org/10.1016/j.giq.2021.101619.
[32] A. Simonetta, M. C. Paoletti, M. Muratore, A new approach for designing of computer architectures using multi-value logic, International Journal on Advanced Science, Engineering and Information Technology 11 (2021) 1440–1446. doi:10.18517/ijaseit.11.4.15778.
[33] D. Steinberg, A. Reid, S. O'Callaghan, F. Lattimore, L. McCalman, T. S. Caetano, Fast fair regression via efficient approximations of mutual information, CoRR abs/2002.06200 (2020). URL: https://arxiv.org/abs/2002.06200.
[34] IBM, Watson openscale, 2022. URL: https://www.ibm.com/it-it/cloud/watson-openscale/drift(Access10-22).