Using the SQuaRE series as a guarantee for GDPR compliance

Alessandro Simonetta
Department of Enterprise Engineering
University of Rome Tor Vergata
Rome, Italy
alessandro.simonetta@gmail.com
ORCID: 0000-0003-2002-9815

Maria Cristina Paoletti
Rome, Italy
mariacristina.paoletti@gmail.com
ORCID: 0000-0001-6850-1184

Alessio Venticinque
Naples, Italy
ORCID: 0000-0003-3286-3137

Abstract—In a context where the availability of information represents the opportunity for companies to gain a competitive advantage in the market through the use of sophisticated AI algorithms, data quality assumes a strategic role. With this paper we want to show that the adoption of an international quality measurement standard such as the one present in the SQuaRE series can on the one hand improve the ethical aspect of machine learning algorithms and on the other hand meet the requirements imposed by the European Community regarding the protection of personal data of citizens in Member States (GDPR). Indeed, although the attention to the protection of personal data is mainly directed towards the aspects of security and confidentiality, in a holistic view we should also evaluate the risks arising from the absence of quality in the data. In this context, we consider consistent and of reference for the international community the choice of the Italian legislator made for the Public Administrations. Since 2013 the Agency for Digital Italy (AgID) has suggested the adoption of ISO/IEC 25012 for public administrations in charge of managing databases of national interest. In the article, we propose a methodological approach that ensures the governance of data quality and some open questions regarding the homogeneity of the selected measures.

Index Terms—ISO 25000, ISO 25012, ISO 25024, SQuaRE series, GDPR, data quality, COVID-19

I. INTRODUCTION

According to The Economist [1], the data represent the new oil for the modern business, not only related to the IT services, but also for what concern the business decision and the marketing campaigns. Many companies are investing in data analysis, machine learning based algorithms and in solutions chosen through data driven approaches. In this scenario they are realizing that the success or their investment are based not only on the amount of data, that however is an important aspect, but mainly on their quality. This could have an impact on the results of machine learning algorithm that are subject to bias on results if the dataset is not properly chosen or have quality problems, i.e. contains unbalanced data. These issues are more evident if the techniques used are taken to extreme as for example in [2] where the use of approximated computing for low power neural network could be more subject to errors. Furthermore, the benefit of using methodologies such as Reinforcement Learning, to contrast the degradation of results and to distribute the decision system as reported in [3] and in

[4], could be frustrated due to poor data quality.

An example is the algorithm used in Florida [5] to score the risk of reiteration for people who went in jail that was subject to bias due to the wrong composition of the data used to trainee it and the features selected. Indeed, the algorithm for calculating the score was trained on a dataset where the criminals were unbalance towards black people and the weight given to the past record of crimes committed and their importance was not properly set. Therefore, where this risk assessment tool of dangerousness and re-offense risk was used African Americans scored higher in criticality compared to Caucasian ones based on the skin color also if their records were less critical.

This case study tell us the importance of the training data set and their quality and the great impact that can have on business decision or citizen life, especially if we concentrate our study on the use that could do public administration or private companies about the people data.

Attention to the use of data, its collection and its quality is a very important issue also in Europe, where for years the legislator has been addressing these issues and investing resources to align regulations with the problems arising from new technologies and new business models [6]. An important step taken in Europe is the introduction of the General Data Protection Regulation (GDPR) 2016/679 [7], defined to harmonize the data privacy laws among the European countries and, in order to remodel the methods and the approaches that the organizations manage the European citizens' data.

The full compliance to this regulation in the past was addressed mainly focused for what concern security issues, but aspect as compliance, integrity and correctness of data are now becoming central. In [8] some issues linked to GDPR are addressed and in particular the compliance of data management and usage for business process. The paper propose three solutions to reduce data maintenance and information loss, avoiding degradation through data minimization during the course of business process. The work covers only part of the accountability principle in that it is not concerned with monitoring and measuring the quality of and maintaining the correctness of the data, but addresses the problem of degradation about information over the time to ensure that

the data once processed is usable for business purposes, in accordance with regulations, even if some of it loses its correctness.

These aspects are taken into account also in [9] that deals with the problems of GDPR compliance in use of Public Geographic Information System for research and practice. The team apply the pseudonymisation to GIS information to guarantee the privacy of the users that share their data.

In [10] the OpenEHR standard is proposed to address some requirements of data privacy regulation. It gives a guideline to guarantee that software for electronic health records are interoperable and secure. The health domain is very sensitive to data quality since the effects of a minimum error can cause irreparable damage with death or serious injuries. However, the work addresses the requirements of integrity and traceability related to data quality; the proposed versioning assure the indelibility of the clinical record preventing any information from being deleted. The creation of a new version of the electronic clinical record is important against the lost, destruction or accidental arm of data. This is only a part of the data quality and a section of the requirements that sensitive information must meet, it does not define KPI or measurement process. Furthermore, the paper is concentrated on clinical data only and miss to consider other information as personal data accuracy that are an important issue. In [11] more details on data quality are presented, and a clear picture of the problem affecting clinical records is reported. However, the study focus only on medical information and is not easy to standardize and applicate to different domains.

According to GDPR, each processing of personal information must be performed in accordance with the quality principles established by art. 5 (adequacy, correctness, update, security, protection, integrity) following the requested criteria of accountability of the data owner. He must guarantee not only the respect of these principles, but also the evidence that he applied all the actions to protect the data (art. 5, par. 2 and art. 24, par. 1). The regulation assigns specific obligation to the responsible for the processing, different compared to those identified for the owner, and in particular the implementation of appropriate technical and organizational measures to ensure the security of the treatment (art. 32) through the concepts of confidentiality, integrity, availability, resilience and ability to restore.

The GDPR is not the only regulation in which quality characteristics (accuracy, completeness, correctness, up-to-date, security, protection, integrity) are reported with a clear meaning but difficult to compare in the absence of a common metric described by a calculation algorithm. Indeed, even the European Solvency II regulation [12], establishes the need for insurance companies to have internal procedures and processes in place to ensure the appropriateness, completeness and accuracy of the data used in the calculation of their technical provisions (art. 82 Data quality and application of approximations, including case-by-case methods, for technical provisions). When granting the basic solvency capital requirement of approval, insurance supervisors must verify the completeness, accuracy,

and adequacy of the data used (art. 104 [12]). In addition, insurance companies must provide for a regular cycle of validation of their internal model that includes assessment of the accuracy, completeness, and adequacy of the data used in the internal model (art. 124 [12]).

In this scenario, where regulations require to ensure certain characteristics related to the management and maintenance of the data over time, many efforts are directed towards ensuring a high level of security for information management, thanks to the application of the ISO/IEC 27000 series [13], but fewer organizations are concerned about managing the risk associated with their management and quality. A good solution could be the application of the ISO 31000 series [14] which provides principles, a framework and a process for managing risk. In section II we describe the state of the art related to the SQuaRE approach to data quality assurance and what benefits are recognized from its adoption in several case studies. A focus will be made on the implications of COVID-19 on the application of GDPR in the health field. In section III we will present the extension of Italian institutions' approach to data quality and we will extend it as a solution to be adopted by private organizations, also as an indispensable support to demonstrate full complaince to GDPR. In section IV we will identify the limitations of this work and how we intend to address them in the future. Finally, in section V we will present concluding remarks.

II. STATE OF ART

A. GDPR Data Quality Compliance in COVID-19 Pandemic

The pandemic emergency boosted the digitalization and the use of online services. The smart working catching on many organizations highlighted problems related to data privacy and data quality, especially for what concern issues that clash with the need for infection tracking. New type of communication and workflows are been developed and adopted during this time with the objective to be accepted by worker and to transmit them trust in data management and security. A central point in this new situation is the compliance of all the data to GDPR. Many organizations had problems related to guarantee its requirements and to manage the information in the right way.

Scientific papers address different aspects of the consequence of COVID-19 pandemic emergency on data privacy and protection in these two years. In [15] a literature review is presented about publications that explore the effect of the COVID-19 outbreak on GDPR compliance. The work identifies some critiques of the regulation and in particular, it focus on the ethical use of health data during the pandemic. Furthermore, the infrastructure use and the absence of controls let the possibility of cross border transfer of data outside the Europe. These aspects are treated also in [16], where the authors study the use of personal information for research activities to defeat the COVID-19 and some criticalities linked to specification of GDPR. The normative has foreseen procedures to support research in pandemic and the processing of sensitive data, included personal and health, but the derogation of some

aspect to national laws is an obstacle to a coordinated global research. This work enhance also the lack of a framework that support the proof of compliance of data management to desired requirements. A particular case of this problem is studied in [17]. The paper describes the problems linked to traceability application offered by mobile devices, Android and iOS based, to monitor the contacts between people and infections. The European legislation is analyzed with respect to sector-specific international rules, as the US Health Insurance Portability and Accountability Act (HIPAA), highlighting the pros and cons of its flexibility in responding to critical health situation.

These scenarios show that the action to protect data and the compliance to regulation is often unfulfilled due to the lack of a common guidelines and the heterogeneity of normative. Although, it is reported that data accuracy is a non-core aspect of data privacy, individuals have the right to correct inaccurate or incomplete personal data that is processed. Using the SQuaRE series as a data quality measurement standard and in order to support GDPR compliance provides for a single reference with respect to individual national regulations helping to harmonize the application of legislative specializations in different countries.

B. The SQuaRE Approach

In this section, in order to support our work, we report an extensive scientific production that focuses on the application of data quality standards and in particular the SQuaRE series to achieve measurable and well-defined goals. Such approaches form the basis of our proposal that aims to achieve full GDPR compliance by achieving the right level of accuracy and satisfying the entire data quality characteristic by applying the SQuaRE series.

Some studies propose a framework for data quality evaluation to let organizations to be able to support and maintain data quality. In [18] the described framework is based on ISO/IEC 25012 [19] and ISO/IEC 25024 [20] and consist of a software that recognize several patterns that identify common failures of organization and help to evaluate the KPI defined for this patterns obtaining a clear picture of the data quality. Small organizations could be not prepared in the application of the standards to their data. The framework presented and the tools, that are starting to spread, represent a solution, that, if widely adopted, can help such companies to guarantee the compliance to the standards with advantage for services results and business decisions.

The problem of data quality and in particular of bias present in dataset used for machine learning algorithm is studied in [21]. This type of systems go under the name of Automated Decision Systems (ADM) as described in [22]. These are tools in which the decision is taken by an AI algorithm in autonomy. The use of algorithms in software involves aspects of daily life such as the marketing campaigns to suggest to customers on a web platform. The work aim is to verify the balance of the dataset before it will be used in prediction algorithm and it cause discriminatory problems. The paper defines metrics with the dual purpose of identifying bias and providing solutions

that can be used within the measurement framework of the serie SQuaRE. The integration of the proposed metrics into the data preparation pipeline for machine learning with the analysis of intrinsic properties of dataset could anticipate the emergence of discriminatory behavior of algorithms that in particular case may contravene laws or infringe human rights. Nowadays, the most successful organizations are those who are able to collect data, select the right set and guarantee the best quality. Their decisions follow a data driven approach and if the basis are wrong, the strategies implemented and the services offered will be affected with negative consequences. Therefore, organizations to be confident with the results of their processing must trust their data. To achieve this level of confidence organizations are implementing the regulation present into the standards and applying process and framework. Many of them are applying data quality evaluation process and data quality management in order to obtain certification for their repositories and not only for the software that they use to process them. In [23] are reported three case studies of data quality evaluation and certification process about repositories. An independent entity verifies and certifies that the company's database complies with the requirements defined in the regulation. The organizations are different each other for what concern dimension and business domain. The two visions are analyzed to evaluate the impact of the adoption of the ISO/IEC 25012, ISO/IEC 25024 and ISO/IEC 25040 [24] and their benefit recognized in the three organization before and after the process. The case studies consist, for each of them, of two phases. The first evaluate the data quality and address the issues found; the second consist in another evaluation on the improved databases to obtain the certification. 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.

An element often found in these articles is related to the use of open formats for archiving aspects to promote portability regardless of the technology used. These are all concepts defined in the SQuaRE series.

III. THE SOLUTION PROPOSED

Previously, we have shown that there are ambiguous interpretations about characteristic of data quality linked to regulations and laws that lead to heterogeneous approaches and uncertainty in the level of adherence to the requirements that must be met. These issues are addressed by scientific works that, also with the support of some case studies, manage to define methodologies and processes to achieve a high level of data quality through the application of standards. In the following we will show our proposal starting from what has been done by the Italian Public Administration to achieve the levels of data quality it needs.

A. Italian Digitization Process

In Italy, the process of modernization and digitization of the Public Administration began with the Digital Administration Code (Legislative Decree n.82 of March 7, 2005), which states that public administration data must be made available and accessible with information and communication technologies that allow their use and reuse.

In 2013, AgID (Digital Italy Agency) with the resolution n. 68 defines, for databases of national interest, the compliance with the quality characteristics defined in the international standard ISO/IEC 25012 "Data quality model".

The adoption of the standard for databases of national interest is thus the reference for both the Public Administration and private companies, also in order to enhance the value of information assets and improve the quality of services offered and their efficiency. The adoption of this data quality governance system responds to their growing need for dissemination and transparency of the information processed. At the beginning, the AgID, in order to simplify the adoption of ISO/IEC 25012, had identified a minimum set of characteristics (accuracy, completeness, consistency and currentness) from which to start and then extend to the entire set. Moreover, in the following years through the Three-Year Plan for Public Administration, an essential tool to promote the digital transformation of the Italian Public Administration, the importance of the use of ISO/IEC 25012 is reaffirmed.

In particular, in the Plan for the three-year period 2020-2022 [25] AgID promotes the increase of data and metadata quality (OB.2.2) of which the increase in the number of open datasets conforming to a subset of quality characteristics derived from the ISO/IEC standard (R.A.2.2b).

B. Method

The proposed approach is based on four main phases (Fig. 1): the initial design, the exercise of the measurement process, the evaluation of the obtained results, and the last phase about the identification of improvement actions. In the initial design phase we need to choose the reference context (conceptual, logical and physical level), we define the quality requirements and the quality model of data to adopt. During this phase it is evaluated the opportunity to reduce/extend the model (override/overload of characteristics and/or measures) and identify the target entities in the appropriate phase of the data life-cycle. Fig. 2 shows the relationship between the data quality characteristics provided in ISO/IEC 25012 and the GDPR principles, noting that these relationships have been highlighted as examples and depend on the context of application. Since each quality characteristic is calculated through the sum of the contribution of several sub-measures, each organization can choose to give more or less importance to individual contributions by assigning a weight to them.

In addition, if we consider comparing different quality characteristics, we may find a different granularity of the constituent contributions to the measures. In this case, the organization may choose to adopt new sub-characteristics or select a subset of them in order to make the measurement system balanced.
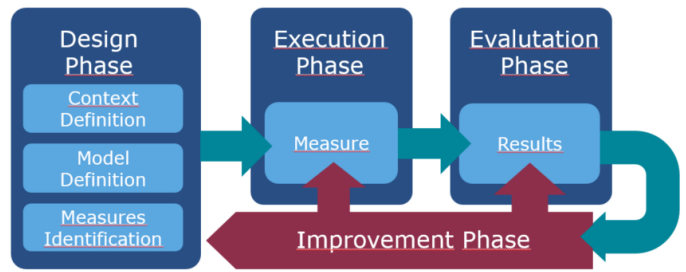

Fig. 1. Data Quality Governance.

The measurement process can take place through commercial software or, more simply, by building a set of modular queries, reusable and moldable on different realities depending on the model identified. The evaluation of the results must be carried out by an expert who assesses the level of quality achieved at the end of each iteration and the progress of the same, at various levels of granularity, over time. The organization has to implement, if needed, improvement actions to mitigate risk and improve services through greater IT efficiency and enhancement of data as an asset.

C. Expected Results

The application of the proposed approach allows the organization to measure the quality of their data and to identify issues related to their acquisition and management. The organization can have the possibility of quality controlled and certified information. The automatization of data management and control eliminates manual queries and cleaning operation. High quality data support the collaboration and the exchange of information between internal and external company structures.

On the other hand, if organizations do not control data quality, they may not only be exposed to bias in sensitive data due to incomplete dataset, but are also subject to different types of risk [26].

The methodology proposed in this paper allows the organization to demonstrate that all activities have been put in place to control and properly manage the data.

The application of a standard capable of supporting the achieving the quality objectives common to the public and private sectors, allows to trace a collaborative direction between these two actors, establishing a virtuous cycle where the application of the same principles makes it easier for both the verify and the prove of compliance with certain national or European rules.

IV. LIMITATION AND FUTURE WORKS

In this paper, are not considered problems linked to the traceability and the validity of the source of information that can play a key role in automate the process of data controls over the time. In a scenario where the organization has traceability on the source of its data or part of it, and can easily know its validity status, it can automate the measurement of its quality and take action to meet its targets.
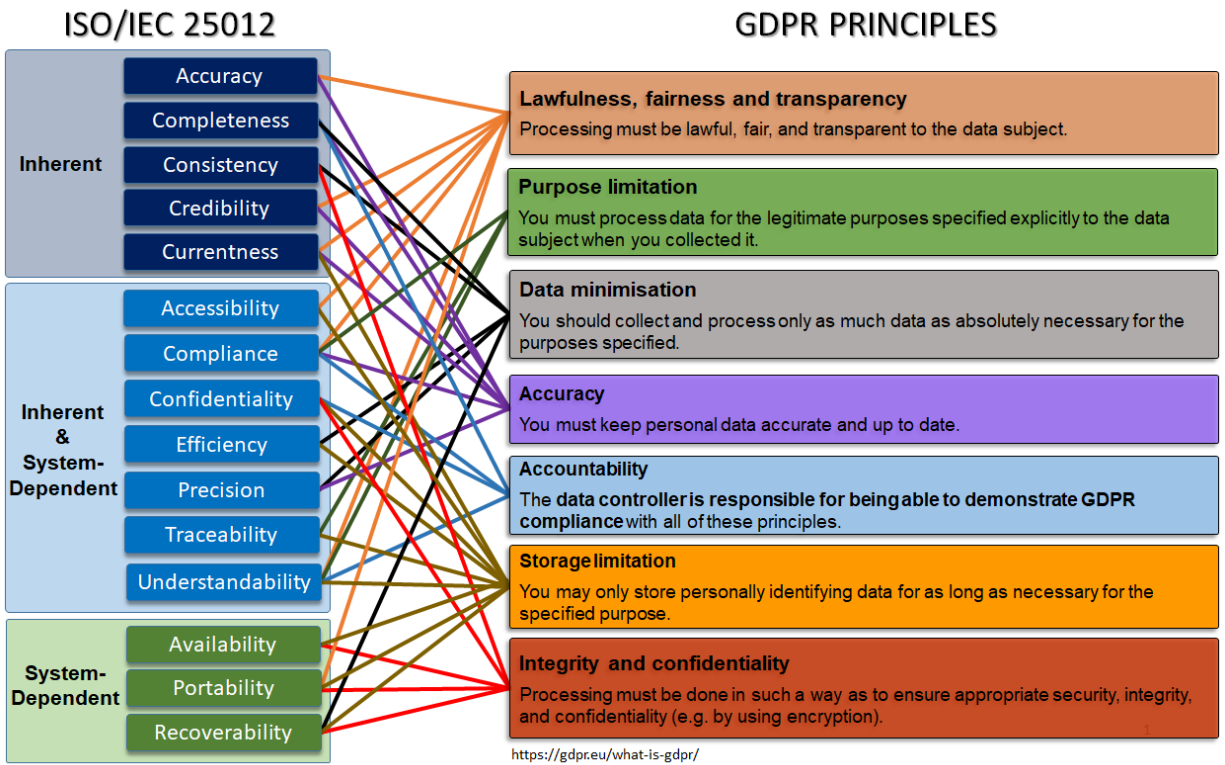
Fig. 2. Relationship between ISO/IEC 25012 characteristics and GDPR principles

As we presented, the data integrity has great importance in those applications where the quality of the services is linked to the composition of dataset. The correctness and the absence of malicious acts is paramount to avoid that organizations can offer a wrong results to their customers or make wrong business decision. The spread of services based on corrupted data could distribute misinformation and compromise compliance with the directives.

In order to get this target we will evaluate in our future work the use of blockchain technology. Its use is starting to be analyzed especially for domain as healthcare and food and beverage as reported in [27]. In this approach, the data are distributed over a network of nodes that approve the correct transaction, with a consensus algorithm, and reject the malicious one. This architecture guarantee the absence of a single point of failure and of central control that could be a point of attack. The data are stored in chain of blocks, where each of this contains the data, a timestamp and the hash of previous block. If the data inside a block is changed, its hash will be different from the hash stored into the chain, so the blocks will be invalidate.

These features are what we need to guarantee the integrity and the validation of data and between the different methodologies that implements the blockchain we will consider the use of the Ethereum and his smart contracts functionality. These smart contracts enforce a contract or an agreement between parties through code without the use of an authority.

The Ethereum Virtual Machine (EVM) provide the services to publish the smart contracts on the Ethereum blockchain. These can be used to store variables within them, which in our case can be the data or information related to them. If on the one hand this allows to validate the contents in the blocks through automatic operations, public and shared, there is a limitation in the use of the blockchain due to the size of the transactions because of the content that you want to put in the blocks. In particular, it is difficult to ensure that every participant agrees to and complies with the relevant rules on personal data protection in public blockchains. Aspects such as a data principal, a data fiduciary, or a data processor on a blockchain network have no clear demarcation.The compliance of these aspects with the GDPR and how the SQuaRE series can help to solve them will be discussed in more detail in future works due to their complexity of treatment.

V. Conclusion

GDPR compliance is well addressed by following compliance with the ISO/IEC 27000 and ISO/IEC 31000 series, however the lack of quality in some features mentioned in art. 5 of the GDPR can lead to errors that impact European citizens. For example, a health recall campaign for a cancer prevention screening sent to an outdated residential address causes harm to the citizen who is not reached by the communication. The use of blockchain technology can help manage this type of

situation: when the organization that owns the data (i.e., the address book) needs to change information in the blockchain, it must run a consensus algorithm with the other parties, so all actors are informed of the change. Also, if the data is changed without applying the rules of the blockchain, everyone else knows, automatically, that that data is invalid. The use of such technology in applications of this type is not yet mature, and in-depth studies on the choice of appropriate algorithms to ensure compliance with regulations must be pursued. The ISO/IEC 25000 series makes it possible to avoid this type of problem by continuously measuring the quality of the data held by organizations. Moreover, the scientific literature is full of examples where the absence of quality in the learning data of an automated decision making system leads to biased analyses especially when the sensitive attributes describing the individuals in the knowledge base are incomplete. Some studies [28] are trying to relate the unbalance of the learning dataset with respect to the fairness of automated classifications. Thus, the importance of data quality is becoming a strategic goal for many companies that often find themselves using replicas of out-of-date data.

REFERENCES

[1] The Economist, "The world's most valuable resource is no longer oil, but data," *The Economist, USA*, 6th May 2019.

[2] G. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Matta, and M. S. S. Patetta, M.and Re, "Approximated computing for low power neural networks," *TELKOMNIKA*, vol. 7, pp. 1236–1241, 2019.

[3] M. Matta, G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, A. Nannarelli, M. Re, and S. Spanò, "A reinforcement learning-based qam/psk symbol synchronizer," *IEEE Access*, vol. 7, pp. 124 147–124 157, 2019.

[4] L. Canese, G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Re, and S. Spanò, "Multi-agent reinforcement learning: A review of challenges and applications," *Applied Sciences*, vol. 11, no. 11, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/11/4948

[5] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, "Machine bias : There's software used across the country to predict future criminals. and it's biased against blacks." *https://www.propublica.org/*, 2016.

[6] Council of Europe, "Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems," 2020.

[7] 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.

[8] R. Zaman and M. Hassani, "On enabling gdpr compliance in business processes through data-driven solutions," *SN Computer Science*, vol. 1, no. 4, pp. 1–15, Jun. 2020.

[9] K. Hasanzadeh, A. Kajosaari, D. Häggman, and M. Kyttä, "A context sensitive approach to anonymizing public participation gis data: From development to the assessment of anonymization effects on data quality," *Computers, Environment and Urban Systems*, vol. 83, p. 101513, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0198971520302465

[10] M. Sousa, D. Gonçalves-Ferreira, C. Pereira, G. Bacelar, S. Frade, O. Pestana, and R. Cruz-Correia, "openehr based systems and the general data protection regulation (gdpr)," *Studies in health technology and informatics*, vol. 247, pp. 91–95, 01 2018.

[11] 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, and D. I. Fotiadis, "Medical data quality assessment: On the development of an automated framework for medical data curation," *Computers in Biology and Medicine*, vol. 107, pp. 270–283, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482519300733

[12] European Parliament and Council of Europe, "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.

[13] *International Organization for Standardization, "ISO/IEC 27000:2018, "Information Technology, Security Techniques, Information Security Management Systems,Overview and Vocabulary"*, International Organization for Standardization Std., 2018. [Online]. Available: https://www.iso.org/standard/73906.html (accessed Nov, 2021)

[14] *International Organization for Standardization, "ISO 31000:2018(en) Risk management — Guidelines"*, International Organization for Standardization Std., 2018. [Online]. Available: https://www.iso.org/iso-31000-risk-management.html (accessed Nov, 2021)

[15] S. McLennan, L. Celi, and A. Buyx, "Covid-19:putting the general data protection regulation to thetest," *JMIR Public Health and Surveillance*, vol. 6, 04 2020.

[16] R. Becker, A. Thorogood, J. Ordish, and M. J. Beauvais, "Covid-19 research: Navigating the european general data protection regulation," *J Med Internet Res*, vol. 22, 2020. [Online]. Available: http://www.jmir.org/2020/8/e19799/

[17] L. Bradford, M. Aboy, and K. Liddell, "Covid-19 contact tracing apps: A stress test for privacy, the gdpr, and data protection regimes," *Journal of law and the biosciences*, vol. 7, p. lsaa034, 05 2020.

[18] J. Calabrese, S. Esponda, and P. M. Pesado, "Framework for Data Quality Evaluation Based on ISO/IEC 25012 and ISO/IEC 25024," in *VIII Conference on Cloud Computing, Big Data & Emerging Topics*, 2020. [Online]. Available: http://sedici.unlp.edu.ar/handle/10915/104778

[19] *International Organization for Standardization, "ISO/IEC 25012:2008 Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model"*, International Organization for Standardization Std., 2008. [Online]. Available: https://www.iso.org/standard/35736.html (accessed Nov, 2021)

[20] *International Organization for Standardization, "ISO/IEC 25024:2015 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of data quality"*, International Organization for Standardization Std., 2015. [Online]. Available: https://www.iso.org/standard/35749.html (accessed Nov, 2021)

[21] A. Simonetta, A. Trenta, M. C. Paoletti, and A. Vetrò, "Metrics for identifying bias in datasets," *SYSTEM*, in press.

[22] R. Fahy and N. Appelman, *Netherlands/Research*, 2020, pp. 164–175, chapter in: Report Automating Society 2020, Chiusi, F., Fischer, S., Kayser-Bril, N. & Spielkamp, M. eds., Berlin: AlgorithmWatch, October 2020. [Online]. Available: https://www.ivir.nl/publicaties/download/Automating-Society-Report-2020.pdf/

[23] F. Gualo, M. Rodriguez, J. Verdugo, I. Caballero, and M. Piattini, "Data quality certification using iso/iec 25012: Industrial experiences," *Journal of Systems and Software*, vol. 176, p. 110938, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0164121221000352

[24] *International Organization for Standardization, "ISO/IEC 25040:2011 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Evaluation process"*, International Organization for Standardization Std., 2011. [Online]. Available: https://www.iso.org/standard/35765.html (accessed Nov, 2021)

[25] Agenzia per l'Italia Digitale, *Piano Triennale 2020-2022*. Presidenza del Consiglio dei Ministri, 2020.

[26] A. Simonetta, A. Vetrò, M. C. Paoletti, and M. Torchiano, "Integrating SQuaRE data quality model with ISO 31000 risk management to measure and mitigate software bias," *IWESQ 2021*, in press.

[27] J. Grecuccio, E. Giusto, F. Fiori, and M. Rebaudengo, "combining blockchain and iot: Food-chain traceability and beyond."

[28] A. Vetrò, M. Torchiano, and M. Mecati, "A data quality approach to the identification of discrimination risk in automated decision making systems," *Government Information Quarterly*, vol. 38, no. 4, p. 101619, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0740624X21000551