Trust Bridging Data Governance and Open Science Adoption in Higher Education Institution

Muhamad Faizal Zulkifli[1], **Norizah Mustamil**[1*], **Zarina Zakaria**[1] and **Adida Md Amin**[2]
[1]Faculty of Business & Economics, Universiti Malaya, Kuala Lumpur, MALAYSIA
[2]Law Library, Universiti Malaya, Kuala Lumpur, MALAYSIA
e-mail: norizahmm@um.edu.my

ABSTRACT
In the current advancement of technology and digitalization era, vast amounts of data has been produced every second, which not only advances scientific research but also plays an important role in facilitating socio-technical development and sustainable development agenda. The usage of open science has positive images and perceptions among students and scientists, which open science publications are more favorable than non-open science publications. Adoption of open science has been proposed as a strategy for universities in developing countries to land in higher rankings and provides opportunities for the public to voluntarily participate voluntarily in the research process. However, not all data is suitable to be shared publicly due to privacy, ownership, trust, and incentive, these concerns need to be solved in enticing higher adoption of open science. Adopting open data or open science is still a challenge especially in gaining trust, which is the prerequisite of open science. To address this shortcoming, in this perspective, we conceptualize the data governance components to gain trust in adopting open data among academicians in the context of higher education institutions. Drawing from data governance taxonomy, this paper posits that a) data quality, b) data security, c) data architecture, d) metadata, e) data lifecycle, f) data storage, and g) data pricing has significant influence towards trust and open science adoption. This paper also argues that trust mediates the relationships of a) data quality, b) data security, c) data architecture, d) metadata, e) data lifecycle, f) data storage, and g) data pricing on open science adoption. Such understanding of data governance and trust in adopting open science would enable a more effective strategy for a better adoption rate of open science in higher education institutions.

Keywords: Open Science, Data Governance, Trust, Higher Education Institution

INTRODUCTION

In the current world that has advanced technology in knowledge discoveries has resulted in tonnes of data to be generated every seconds, with the abundance of data and knowledge has helped us not just in shaping the world but also saving humankind such as during COVID19, where the open science has helped all scientist around the world to communicate and sharing data in finding the solutions to pandemic like COVID19 (Viseur, 2021). Open science is not just advancing scientific research but also plays an important role in facilitating socio-technical development and sustainable development agenda (Samuel & Lucivero, 2020). In the context of higher education institutions, the usage of open science is growing but still limited, nevertheless the perception of open science among higher education users is positive especially among the students and researchers, which open science publications are more favorable than non-open science publications (Schneider et al., 2022).

To understand the landscape of open science in higher education institutions is very important as higher education institutions are one of the organizations that are active in research. So, having more higher education institutions to participate in open gives higher data richness to the open science databases. The adoption of open science in higher education is getting the attention of scholars (Ivančević & Luković, 2018). In addition, the usage of open science has been identified as one of the key components and strategies in landing the university ranking to the higher rankings (Kurniasih et al., 2018). Open science provides opportunities and platforms for the public and researchers to voluntarily participate in the research process, which would benefit humankind (Abd. Rahman, 2019).

However, not all data is suitable to be shared publicly due to its nature including the privacy, ownership, trust, and incentive (Li et al., 2022). Adopting open data or open science is still a challenge especially in gaining trust of the users, which trust is one of the main prerequisite of open science to be adopted by researchers and change the research landscape (Clark, 2021). Hence, these concerns need to be understood and solved to have higher participation from higher education institution users in open science. The understanding of data governance in open science from the higher education institutions perspective is crucial to gain trust among researchers to adopt open science (Demchenko & Stoy, 2021).

To address this shortcoming, the main objective of this paper is to understand how data governance develops trust among higher education institution users to adopt open science. In order to understand the main objective, drawing from data governance taxonomy (Abraham et al., 2023), this paper posits that a) data quality, b) data security, c) data architecture, d) metadata, e) data lifecycle, f) data storage, and g) data pricing has significant influence towards trust and open science adoption. This paper also argues that trust mediates the relationships of a) data quality, b) data security, c) data architecture, d) metadata, e) data lifecycle, f) data storage, and g) data pricing on open science adoption. The understanding of data governance and trust in adopting open science would enable a more effective strategy for a better adoption rate of open science in higher education institutions.

LITERATURE REVIEW

Open Science

Over the past decades, the advancement of technology and the digital world has evolved the world of science towards transparency, reproducibility, and openness, which has resulted in a movement known as 'Open Science' (Armeni et al., 2021). The challenges of reliability and accountability of scientific works have gaining attention of scholars especially on the transparency of the research, in light of these issues, researchers and scholars are driven to increase the reliability and transparency in various aspects of the work beginning from the initial research until the finish products (Bezjak et al., 2018; Munafò et al., 2017; Nosek et al., 2015; Nosek et al., 2018; Stall et al., 2019).

The concern over accessibility and transparency of scientific research also gaining attention from journals, funders, and policy makers where they urge and having expectation to scholars to increase the accessibility and transparency of their search, finding, and also products (Aczel et al., 2020; Burgelman et al., 2019; Morey et al., 2016). So, by adopting openness in research, scholars would work in a less error environment (Hales et al., 2019), and have higher visibility to peers in the same research area and also to scientist from other discipline, which also resulted in higher citation rates (Colavizza et al., 2020). In addition, participating in open science practices would promote and facilitate the sharing and reuse of data, materials, and code in the scientific community (Allen & Mehler, 2019), which would increase scholars' outputs and literacy, besides it also increase the trust in the scholarly process (Tennant et al., 2016).

By promoting and adopting open science, it is not only giving tangible benefits to individual researchers, but also benefits the scientific community and the society at large (Armeni et al., 2021). But, in many open science occasions, most of the time it often attracts innovators and early adopters only, which creates so called 'open science bubbles'. (Armeni et al., 2021). While, it is good to have the early adopters of open science in doing their research workflows, but a critical number of adoptions are needed especially among mass scientist and scholars so that open science can move from open science advocacy to actual behavior, but this still remains challenging especially when open science is not widely and normatively accepted by the scientific communities yet (Armeni et al., 2021)

As part of transitioning towards open science, it has promoting the need for the research environment to adapt to new societal and technological advancement over the past years (Burgelman et al., 2019), such as the usage of web-based technologies and social media networks as regular tools for data collection, sharing, analysis, and collaboration (Voytek, 2017). So, the abundance of available data and its availability to open science has raised the concerns among scholars on how the data has been managed.

Data governance

Data governance refers to a framework that gives structure and formalization in terms of data management (Morabito, 2015). In order to have a good data governance, organizations need to specify what must be governed such as the scope of the data (Abraham et al., 2019), who are responsible in governing the data, such as the roles and governance bodies (Otto, 2011), and what are the decisions must be made in data-related areas such as the data governance decision domains (Abraham et al., 2019; Lee et al., 2019). This study draws the last component. Based on Abraham et al. (2023), we describe the seven data governance decision domains: (a) data quality; (b) data security; (c) data architecture; (d) metadata; (e) data lifecycle; (f) data storage and infrastructure, and (g) data pricing towards trust in adopting open science among higher education institution users.

i) Data quality
Data quality is defined as the ability of data to fulfill the usage requirement in a specific context (Khatri & Brown, 2010). Data quality is evaluated based on quality dimensions of completeness, credibility, accuracy, timeliness, and consistency of data (Khatri & Brown, 2010). In the scientific literature, preventive and reactive measures are two proposed measures in managing data quality (Otto et al., 2012). In the context of open science, preventive measures would inhibit data providers from onboarding data products with insufficient quality. For example, data providers take additional steps in testing the quality for the data through automated test scripts before putting the data available for consumption (Smith et al., 2016). While, the main aim of reactive measures is to support the identification and reporting of data quality issues after the data has been made available for consumption. The example includes rating systems that allow consumers to rate and provide feedback on the data (Zuiderwijk et al., 2014) or data providers (Ramachandran et al., 2018). The data quality has been backed up by trust in adoption of data in the Internet of Thins (IoT) setting (Byabazaire et al., 2020).

Hence, in the context of open science in higher education settings, we posit that a high standard of data quality will gain the trust of higher education institution users in adopting open science. By having preventive and reactive measures in place for open science, it creates higher data quality, which helps in gaining trust among the users of higher education institutions to adopt open science as part of their research behavior.

Proposition 1: Data quality helps in gaining trust of higher education institution users in adopting open science

ii) Data security

Data security is defined as the preservation of security measurements including the accessibility, authenticity, availability, confidentiality, integrity, privacy, and reliability of data (Carretero et al., 2017). So in the context of data consumption publicly, the concerns of requirement are including the control of when, to whom, and to what extent data is being made available for consumption (Tzianos et al., 2019) and how and where the data being used (Otto & Jarke, 2019). Data security is always associated with trust in adopting any technology (Sun et al., 2013)

In order to establish data confidentiality, data providers use encryption techniques (Tzianos et al., 2019). In order to protect data sensitivity during usage, data providers are adopting methods that fully restrict raw data access and only allow certain part of the data to be accessed, such as the identity data has been hidden using anonymization techniques (Ha et al., 2019). Homomorphic encryption is also being used in enabling mathematical operation on encrypted data (Roman & Stefano, 2016). Data usage terms have been adopted in controlling and protecting data ownership, where these terms will describe the appropriate data usage (Otto & Jarke, 2019; Tzianos et al., 2019). In addition, data terms and contracts help in negotiating and assuring the authorizations, obligations, and prohibitions on data covered by the contract (Allen et al., 2014). This would enable data providers to have a remedy against data consumers in case of contract infringements (Truong et al., 2012).

So, in the context of open science and higher education institutions, we posit that data security helps in gaining trust among higher education institution users to adopt open science in their work environment. This includes the specification of confidential data storage, data access control, confidential data usage, and data usage control. We argue that by having high control of these 4 aspects of data security, it leads to higher trust among higher education institutions to adopt open science in their research.

Proposition 2: High data security will gain trust of higher education institution users in adopting open science

iii) Data architecture

Data architecture is defined as a set of data specifications, which is used as guidelines for data requirements and data integration (Abraham et al., 2023), that consist of comprehensive data models on a conceptual, logical, and physical level (Watson et al., 2004). Data architecture is important in introducing transparent data for consumption (Luciano et al., 2017).

In the context of the data marketplace, data standards are often referred to as being important to supporting interoperability and data exchange between data providers and data consumers (Lis & Otto, 2020). In the data marketplace, data format ranging from standardized through proprietary to hybrid, where data marketplaces define standardized vocabularies and formats, which all participants in the marketplace must follow (Otto & Jarke, 2019), while proprietary approach allows data providers to offer their data products using their own proprietary data formats (Özyilmaz et al., 2018). But, the most convenient approach for both parties, which is data providers and consumers is a hybrid approach, where data providers can supply data in proprietary format, which later will be automatically normalized by the data platform using standardized data model (Nagorny et al., 2018).

So in the context of open science in higher education institutions, we argue that comprehensive data architecture through the right data format such as standardized data format, proprietary data format, or hybrid data format is able to promote data transparency that helps to build trust among higher education institution users in adopting open science.
Proposition 3: Comprehensive data architecture develops trust of higher education institution users in adopting open science

iv) Metadata

Metadata is defined as data about data (Abraham et al., 2023), where metadata is giving meaning and context to data by providing a structured description of the content, quality, and other characteristics of data (Khatri & Brown, 2010). In the data marketplaces context, rich information of metadata is crucial in supporting data consumers especially when finding data of interest (Tzianos et al., 2019), identifying the usefulness of the data (Ramachandran et al., 2018), and precisely interpreting and processing data (Zuiderwijk et al., 2014).

There are two approaches regarding the metadata vocabulary in the scientific literature, which are specific metadata vocabulary and standardized metadata vocabularies (Abraham et al., 2023). Specific metadata data vocabulary is used by data providers to describe and publish metadata, while data consumers use it to look up and retrieve metadata (Otto & Jarke, 2019). A few examples of standardized metadata vocabularies are CERIF and DCAT (Zuiderwijk et al., 2014). The well-established metadata standard would result in data authenticity.

So, in the context of open science of higher educational institutions, established metadata vocabulary (standardized vocabulary or marketplace-specific vocabulary) would enable data providers and users to search relevant and authentic data. The identification of the right and established data architecture helps researchers in searching for the right data which leads to gaining their trust in adopting open science.

Proposition 4: Established metadata vocabulary gains trust of higher education institution users in adopting open science

v) Data lifecycle

The data lifecycle is defined as the whole lifecycle of data starting from collecting, creating, using, maintaining, archiving, and until deleting the data (Khatri & Brown, 2010). For example, in the context of data marketplace, the main life cycles of the data phases are data onboarding, data discovery, data purchase, and data usage (Abraham et al., 2023). Where, during the data onboarding, the data providers would capture, create, and store the data, which later is made available for consumers' consumptions (Otto & Jarke, 2019). During the data discovery phase, the consumers will search the right data based on their goals and consumptions (Ramachandran et al., 2018). In the data purchase phase, consumers would pay in the exchange of the data, and the data providers will give access to the users to access the purchased data (Tzianos et al., 2019). In the final stage, which is data usage stage, the consumers will use the data in achieving their aims or goals such as by enriching and aggregating it (Otto & Jarke, 2019).

In the context of open science at higher education institutions, we posit that the understanding of a suitable data cycle develops trust to the higher education institution users to adopt open science. It is important to understand which data lifecycle is more appealing to researchers at higher education institutions (data trade focus vs data usage focus). The understanding of the right data lifecycle helps in gaining trust among researchers, which leads to higher adoption of open science at higher education institutions.

Proposition 5: Suitable data lifecycle promotes trust of higher education institution users in adopting open science

vi) Data storage and infrastructure

Data storage and infrastructure is information technology (IT) artifacts that are responsible for effective data management (Tallon et al., 2013). How data must be stored is always a question in data management (Abraham et al., 2023). There are three main approaches in data storage which are the centralized, decentralized, and hybrid storage approaches (Spiekermann, 2019).

In centralized approach, data are provided by data provider via a central location such as cloud storage service, while in decentralized approach, data will be stored at data provider facilities, and the hybrid data storage approach is the combination of both the centralized and decentralized approaches (Abraham et al., 2023). The location and storage of data highly influence trust of the users (Dixit et al., 2021).

So, in the context of open science in higher education institutions, we posit that a secured data storage and infrastructure develop trust among the higher education institution users to adopt open science. The understanding of data storage and infrastructure from higher education institutions helps to gain their trust in adopting open science.

Proposition 6: Secured data storage and infrastructure develops trust of higher education institution users in adopting open science

vii) Data pricing

When the exchange of data involves various parties, the question that arises is how to price data relevantly (Abraham et al., 2023). In the data marketplace, there are main pricing models that have been used, which are pay-per-use and subscription-based pricing models based on their business models. Through a pay-per-use model, the data marketplace would charge consumers based on the data consumption (Spiekermann, 2019; Truong et al., 2012). While via subscription based pricing strategy, consumers will be granted access to data for a certain period of time. Other than these two pricing models, data can be provided free of charge whenever allowed by a data provider, which is normally done by public authorities and non-profit organizations (Spiekermann, 2019). There is also hybrid pricing strategy being used by data providers, where basic data is supplied free of charge but providers are charging premium prices for detailed data (Thomas & Leiponen, 2016). In addition, data pricing would enable the right price for data (Truong et al., 2012). In the data marketplace, other than fixed prices, they also adopting more dynamic pricing such as bidding (Parra-Arnau, 2018), progressive pricing (Spiekermann, 2019), the “pay what you want” approach (Zuiderwijk et al., 2014), and packaged pricing (Spiekermann, 2019).

So, in the context of open science in higher education institutions, we posit that reasonable and sustainable data pricing will give trust among higher education institutions to be the data provider and data user of open science. It is important to understand which data pricing strategy is more suitable for higher education institution strategy, as the right pricing strategy helps in developing trust in adopting open science.

Proposition 7: Suitable and sustainable pricing develops trust among higher education institution users in adopting open science

DISCUSSION

Higher education institutions are one of the parties that are actively involved in vast amounts of research that generates tonnes of data, which the data can be used and transformed to various outputs that can be beneficial to various parties such as researchers, scholars, businesses, policy makers and others. The movement of the scientific community towards transparent science such as open science is a noble movement in advancing the state of knowledge. But, the question of how this data is being governed is always a concern not just to the data users, but also the data providers.

Hence, this paper posits that data governance as the main factor to gain trust of higher education institution users in adopting open science. There are seven main components that would influence trust of higher education institution users to adopt open science, which are : (a) data quality; (b) data security; (c) data architecture; (d) metadata; (e) data lifecycle; (f) data storage and infrastructure, and (g) data pricing. Establishing good data governance in the context of open science will give more confidence among researchers and users in higher education institutions to adopt open science as part of their research cycle.

This paper also provides theoretical advancement for trust literature, which is by examining the data governance taxonomy towards open science via trust, it gives understanding on how data governance develops trust among higher education institution users to adopt open science. Gaining trust is one of the most challenging in adoption behavior, so the understanding of data governance influence will provide the fundamental understanding to the researchers.

CONCLUSION

The provided propositions help researchers to understand how data governance dimensions (a) data quality; (b) data security; (c) data architecture; (d) metadata; (e) data lifecycle; (f) data storage and infrastructure, and (g) data pricing influence trust among higher education institution users to adopt open science in their research cycle. The identification of how each dimension of data governance influences trust in adopting open science among higher education institutions will give insight to the data owners and data providers to facilitate the needs of higher education institution users in adopting open science. We conclude that high data governance practice by data owners and providers will gain trust among higher education institution users to adopt open science in their research.

REFERENCES

Abd Rahman, N. 2019. The Need for Open Science. *Journal of Research Management & Governance*, 2(1), 22-30.

Abraham, R., Schneider, J., & Vom Brocke, J. 2019. Data governance: A conceptual framework, structured review, and research agenda. *International journal of information management*, *49*, 424-438.

Abraham, R., Schneider, J., & vom Brocke, J. 2023. A taxonomy of data governance decision domains in data marketplaces. *Electronic Markets*, *33*(1), 22.

Aczel, B., Szaszi, B., Sarafoglou, A., Kekecs, Z., Kucharský, Š., Benjamin, D., Chambers, C. D., Fisher, A., Gelman, A., & Gernsbacher, M. A. 2020. A consensus-based transparency checklist. *Nature human behaviour*, *4*(1), 4-6.

Allen, C., Des Jardins, T. R., Heider, A., Lyman, K. A., McWilliams, L., Rein, A. L., Schachter, A. A., Singh, R., Sorondo, B., & Topper, J. 2014. Data governance and data sharing agreements for community-wide health information exchange: lessons from the beacon communities. *EGEMS*, *2*(1).

Allen, C., & Mehler, D. M. 2019. Correction: Open science challenges, benefits and tips in early career and beyond. *PLoS biology*, *17*(12), e3000587.

Armeni, K., Brinkman, L., Carlsson, R., Eerland, A., Fijten, R., Fondberg, R., Heininga, V. E., Heunis, S., Koh, W. Q., & Masselink, M. 2021. Towards wide-scale adoption of open science practices: The role of open science communities. *Science and Public Policy*, *48*(5), 605-611.

Bezjak, S., Clyburne-Sherin, A., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., & Psomopoulos, F. 2018. The open science training handbook.

Burgelman, J.-C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schouppe, M. 2019. Open science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Frontiers in big data*, *2*, 43.

Byabazaire, J., O'Hare, G., & Delaney, D. 2020. Data quality and trust: A perception from shared data in IoT. 2020 IEEE International Conference on Communications Workshops (ICC Workshops),

Carretero, A. G., Gualo, F., Caballero, I., & Piattini, M. 2017. MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000. *Computer Standards & Interfaces*, *54*, 139-151.

Clark, J. 2021. Open science-a question of trust. *Data Intelligence*, *3*(1), 64-70. https://doi.org/10.1162/dint_a_00078

Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., & McGillivray, B. 2020. The citation advantage of linking publications to research data. *PloS one*, *15*(4), e0230416.

Demchenko, Y., & Stoy, L. 2021. Research Data Management and Data Stewardship Competences in University Curriculum. 2021 IEEE Global Engineering Education Conference (EDUCON),

Dixit, A., Singh, A., Rahulamathavan, Y., & Rajarajan, M. 2021. Fast data: A fair, secure and trusted decentralized IoT data marketplace enabled by blockchain. *IEEE Internet of Things Journal*.

Ha, M., Kwon, S., Lee, Y. J., Shim, Y., & Kim, J. 2019. Where WTS meets WTB: A blockchain-based marketplace for digital me to trade users' private data. *Pervasive and Mobile Computing*, *59*, 101078.

Hales, A. H., Wesselmann, E. D., & Hilgard, J. 2019. Improving psychological science through transparency and openness: An overview. *Perspectives on Behavior Science*, *42*, 13-31.

Ivančević, V., & Luković, I. 2018. National university rankings based on open data: a case study from Serbia. *Procedia Computer Science*, *126*, 1516-1525.

Khatri, V., & Brown, C. V. 2010. Designing data governance. *Communications of the ACM*, *53*(1), 148-152.

Kurniasih, N., Hasyim, C., Wulandari, A., Setiawan, M. I., & Ahmar, A. S. 2018. Comparative Case Studies on Indonesian Higher Education Rankings. Journal of Physics: Conference Series,

Lee, S. U., Zhu, L., & Jeffery, R. 2019. Data governance decisions for platform ecosystems.

Li, Z., Mao, F., & Wu, C. (2022). Can we share models if sharing data is not an option? [Review]. *Patterns*, *3*(11), Article 100603. https://doi.org/10.1016/j.patter.2022.100603

Lis, D., & Otto, B. 2020. Data governance in data ecosystems–Insights from organizations.

Luciano, J. S., Sayogo, D., Ran, W., DePaula, N., Jarman, H., Tayi, G., Zhang, J., Hrdinova, J., Pardo, T., & Andersen, D. L. 2017. Building a certification and inspection data infrastructure to promote transparent markets. *International Journal of Electronic Government Research (IJEGR)*, *13*(4), 53-75.

Morabito, V. 2015. Big data and analytics. *Strategic and organisational impacts*.

Morey, R. D., Chambers, C. D., Etchells, P. J., Harris, C. R., Hoekstra, R., Lakens, D., Lewandowsky, S., Morey, C. C., Newman, D. P., & Schönbrodt, F. D. 2016. The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review. *Royal Society open science*, *3*(1), 150547.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. 2017. A manifesto for reproducible science. *Nature human behaviour*, *1*(1), 1-9.

Nagorny, K., Scholze, S., Ruhl, M., & Colombo, A. W. 2018. Semantical support for a CPS data marketplace to prepare Big Data analytics in smart manufacturing environments. 2018 IEEE Industrial Cyber-Physical Systems (ICPS),

Nosek, B., Alter, G., Banks, G., Borsboom, D., Bowman, S., & Breckler, S. 2015. Promoting an open research culture. *Science*, *348*(6242), 1422-1425.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. 2018. The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600-2606.

Otto, B. 2011. Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, *29*(1), 3.

Otto, B., Hüner, K. M., & Österle, H. 2012. Toward a functional reference model for master data quality management. *Information Systems and e-Business Management*, *10*, 395-425.

Otto, B., & Jarke, M. 2019. Designing a multi-sided data platform: findings from the International Data Spaces case. *Electronic Markets*, *29*(4), 561-580.

Özyilmaz, K. R., Doğan, M., & Yurdakul, A. 2018. IDMoB: IoT data marketplace on blockchain. 2018 crypto valley conference on blockchain technology (CVCBT),

Parra-Arnau, J. 2018. Optimized, direct sale of privacy in personal data marketplaces. *Information Sciences*, *424*, 354-384.

Ramachandran, G. S., Radhakrishnan, R., & Krishnamachari, B. 2018. Towards a decentralized data marketplace for smart cities. 2018 IEEE International Smart Cities Conference (ISC2),

Roman, D., & Stefano, G. 2016. Towards a reference architecture for trusted data marketplaces: The credit scoring perspective. 2016 2nd International Conference on Open and Big Data (OBD),

Samuel, G., & Lucivero, F. 2020. Responsible open science: Moving towards an ethics of environmental sustainability. *Publications*, *8*(4), 54.

Schneider, J., Rosman, T., Kelava, A., & Merk, S. 2022. Do Open-Science Badges Increase Trust in Scientists Among Undergraduates, Scientists, and the Public? [Article]. *Psychological Science*, *33*(9), 1588-1604. https://doi.org/10.1177/09567976221097499

Smith, G., Ofe, H. A., & Sandberg, J. 2016. Digital service innovation from open data: exploring the value proposition of an open data marketplace. 2016 49th Hawaii international conference on system sciences (HICSS),

Spiekermann, M. 2019. Data marketplaces: Trends and monetisation of data goods. *Intereconomics*, *54*(4), 208-216.

Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Lehnert, K., Nosek, B., Parsons, M., Robinson, E., & Wyborn, L. 2019. Make scientific data FAIR. *Nature*, *570*(7759), 27-29.

Sun, Y., Chen, X., & Du, X. 2013. Trust analysis method of process behavior for data security exchange. Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC),

Tallon, P. P., Ramirez, R. V., & Short, J. E. 2013. The information artifact in IT governance: Toward a theory of information governance. *Journal of Management Information Systems*, *30*(3), 141-178.

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. 2016. The academic, economic and societal impacts of Open Access: an evidence-based review. *F1000Research*, *5*.

Thomas, L. D., & Leiponen, A. 2016. Big data commercialization. *IEEE Engineering Management Review*, *44*(2), 74-90.

Truong, H.-L., Comerio, M., De Paoli, F., Gangadharan, G., & Dustdar, S. 2012. Data contracts for cloud-based data marketplaces. *International Journal of Computational Science and Engineering*, *7*(4), 280-295.

Tzianos, P., Pipelidis, G., & Tsiamitros, N. 2019. Hermes: An open and transparent marketplace for IoT sensor data over distributed ledgers. 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC),

Viseur, R. 2021. Open data in digital strategies against COVID-19: the case of Belgium. Proceedings of the 17th International Symposium on Open Collaboration,

Voytek, B. 2017. Social media, open science, and data science are inextricably linked. *Neuron*, *96*(6), 1219-1222.

Watson, H. J., Fuller, C., & Ariyachandra, T. 2004. Data warehouse governance: best practices at Blue Cross and Blue Shield of North Carolina. *Decision Support Systems*, *38*(3), 435-450.

Zuiderwijk, A., Loukis, E., Alexopoulos, C., Janssen, M., & Jeffery, K. 2014. Elements for the development of an open data marketplace. Conference for E-Democracy and Open Government,