2023 European Conference on Computing in Construction
40th International CIB W78 Conference
Heraklion, Crete, Greece
July 10-12, 2023

A CARBON DATA TRUSTWORTHINESS FRAMEWORK FOR CONSTRUCTION SECTOR

Jinying Xu[1], Kristen MacAskill[1]
[1] Department of Engineering, University of Cambridge, Cambridge, UK

Abstract

The construction and building sector contributes 37% of global emissions; yet fragmented, manual and untransparent carbon data collection methods create data trustworthiness issues for making informed design and construction decisions. This article reviews practical carbon standards, data models, and data trustworthiness studies. Then, a carbon data trustworthiness framework is proposed. The framework presents attributes associated with data collection and management that influence the trustworthiness of project carbon data. These are collated under four pillars: data availability, data quality, data compatibility, and data security. The critical technical solutions that contribute to trustworthy data are summarized. This framework is a conceptual guidance for standardized carbon data model. It will lay a cornerstone for digitalized and automated carbon management and thus contribute to carbon reduction.

Introduction

The latest Global Status Report for Buildings and Construction (UNEP, 2022) notes that the buildings and construction sector contributed to around 37% of global CO2 emissions in 2021. Local current situations vary, for example, China's construction sector accounts for 35-50% of its national total carbon emissions (Zhu et al, 2022).The buildings and construction sector "remains off track to achieve decarbonization by 2050" (UNEP, 2022, p32); there is no doubt the sector has work to do to decarbonize.

To reduce the carbon emissions from the construction sector, different levels of government across countries are developing regulations and guidance. Some examples are the Act (2021:787) on climate declarations for buildings in Sweden, CAM (Minimum Environmental Criteria) Green public procurement Law (56/2017) in Italy, and EPD mandatory for environmental claims in France (One Click LCA, 2022). A very specific guidance is the UK *PAS* (Publicly Available Specification) *2080:2023 - Carbon management in buildings and infrastructure*. The PAS 2080:2023, while not a formal law, is a key guidance for the infrastructure and construction sector in its carbon management across the "value chain", which is defined as "organizations and stakeholders involved in creating, operating and managing assets and/or networks" (BSI, 2023). It is gaining traction as sector stakeholders are increasingly realizing their obligations under climate-related legislation. Through providing general principles, it outlines the processes of carbon management and the allocation of responsibility between value chain stakeholders. It does not specify carbon management methods and sources of emission factors. It is complemented by a guidance document on carbon management process, implementation actions and responsibilities, with some case studies. There is no specific guidance on how carbon data are collected, reported, shared, and managed.

Various value chain stakeholders are using their own data collection systems at different levels of sophistication, the interaction between organizations can create headaches for those involved. It is observed that the typology of resources and activities in many in-house developed carbon accounting and reporting tools are ambiguous and not comprehensive. They lack requirements of data accuracy, reliability, and accountability. Based on our observations of the transport sector in the UK, carbon data management is clearly not yet a mature practice. There has been some progress in standardizing emissions factors, but significant reliance on manual efforts to record data into spreadsheets. Without an established whole process carbon data management methodology, the current carbon accounting and reporting practices lead to problems in data trustworthiness, barriers in automating carbon accounting, and can lead to a laborious workload.

Under the urgent net zero (or carbon reduction) targets being set around the world, governance and management of carbon data requires appropriate carbon baseline setting, robust carbon accounting, transparent data reporting to track progress, which all contribute to improving the ability to make informed decisions to reduce emissions. However, obtaining reliable data remains a challenge due to poor data collection methods, data silos, and a lack of standardization in data reporting. A clear definition of carbon data trustworthiness and identification of approaches to achieve it is therefore significant. This article aims to clarify the terminology relating to data trustworthiness, and develop a framework, its key pillars and supporting attributes, and technologies for improving carbon data trustworthiness to set a good steppingstone towards improved carbon management.

Existing carbon standards, guidance, and models

'Carbon management' is "assessment, reduction and removal of greenhouse gas emissions during the planning, optioneering, design, delivery, operation, use, end of life (and beyond) of new, or the management of existing, assets, networks and/or systems" (BSI, 2023). This paper is focused on carbon *data* management – i.e., how data is managed to achieve effective carbon management.

Carbon standards and guidance

International organizations, different countries and sectors are all working toward developing carbon guideline and frameworks. To manage the scope of this study, Table 1 focuses on some leading examples identified from mainland Europe and the UK. There are relevant legislations developed by governments, standards developed by the international standard organization (ISO), European standards organizations, and the British Standards Institution (BSI), as well as guidance developed by governments and professional organizations.

Table 1. Selected carbon standards and guideline in the construction sector

Code	Name	Type
ISO 21930 : 2017	Sustainability in buildings and civil engineering works – Core rules for environmental product declarations of construction products and services	International standard
ISO 14067 : 2018	Greenhouse gases – Carbon footprint of products – Requirements and guidelines for quantification	
EN 15978 : 2011	Sustainability of construction works – Assessment of environmental performance of buildings – Calculation method	European standards (ENs)
EN 15804 : 2012	Sustainability of construction works – Environmental product declarations – Core rules for the product category of construction products	
EN 17472 : 2022	Sustainability of construction works – Sustainability assessment of civil engineering works – Calculation methods	
--	Building Regulation Part Z – Whole life carbon (revision proposed in 2022)	UK legislation
--	Promoting Net Zero Carbon and Sustainability in Construction Guidance Note (2022)	UK Government Guidance
Act 2021: 787	Act (2021:787) on Climate Declaration for Buildings	Sweden legislation
Law 56/20 17	CAM Green public procurement Law	Italy legislation
--	MMG: environmental profile of building elements	Belgium building regulations
PAS 2080: 2023	PAS 2080: *Carbon management in buildings and infrastructure*	British Standards Institution
TM65	Embodied carbon in building services: A calculation methodology (2021)	CIBSE guide
--	Net Zero Whole Life Carbon Roadmap (2021)	UK GBC roadmap
--	Professional Statement on the Whole Life Carbon Assessment (2017)	RICS professional standards and guidance
--	IPA Best Practice in Benchmarking 2019	UK IPA guidance

These standards and guidance provide a guiding framework for carbon measurement and management. They do not tend to specify a particular methodology for carbon emissions quantification across the full life-cycle of buildings or infrastructure. For example, the methodology in TM65 only accounts embodied carbon but not construction activity carbon. According to PAS 2080:2016, there are three carbon emission quantification methodologies: calculation-based life-cycle assessment (LCA) calculation-based input-output analysis (IOA), and a measurement-based method. In the calculation-based methods, a rate of activity is combined with an emissions factor for the carbon emissions of that activity; while a measurement-based method measures the physical emissions as it occurs (where the confidence level is dependent on the standards and type of measurement undertaken). LCA is the most applied method where the emissions factor is determined by analyzing the process and activities of a study system, working towards a system boundary (analysis cut-off point) in a bottom-up way. IOA is a top-down method where activity emission factors are determined based on very broad boundaries, based on interconnected economic sector information, and macro (e.g., national, regional or sector) emission factors data. In calculating carbon, every material or activity quantity is multiplied by a corresponding emissions factor. These are summed to find a total number for reporting, but not sufficient for in-depth analysis and optimization at a detailed process level.

Carbon tools

Various carbon data management tools are developed for practical use. Generally, a widely used tool by companies across the world is the GHG emission calculation tool developed under the leadership of the World Resource Institute (WRI). The GHG emission calculation tool is a spreadsheet-based tool with default and customizable emission factors that vary by country or even location (region/city). It calculates the emissions based on three scopes as summarized in Table 2.

Table 2. The three scopes of GHG emissions

Scope	Activity Type
Scope 1	Stationary combustion
	Mobile combustion
	Fugitive emissions from air-conditioning
	Other fugitive or process emissions
Scope 2	Purchased electricity - location based
	Purchased electricity - market based
	Purchased heat and steam
	Scope 2 - Location based + heat and steam
	Scope 2 - market based + heat and steam
Scope 3	Purchased goods and services
	Capital goods
	Fuel-and energy-related activities (not included in Scope 1 or scope 2)
	Upstream transportation and distribution
	Waste generated in operations
	Business travel
	Employee commuting
	Upstream leased assets
	Downstream transportation and distribution
	Processing of sold products
	Use of sold products
	End-of-life treatment of sold products
	Downstream leased assets
	Franchises
	Investments

source: https://ghgprotocol.org/calculation-tools

There are many sector-specific, practical carbon management tools in the UK, such as RSSB (Rail Safety and Standards Board) Rail Carbon Tool, UKWIR (UK Water Industry Research) Carbon Accounting Workbook, and National Highways Carbon Calculator, and Built Environment Carbon Database (BECD).

The RSSB Rail Carbon Tool calculates and analyses the carbon footprints of UK rail projects and activities, identifies and assesses alternative low carbon options, selects low carbon solutions, allows for building information modelling (BIM) integration, and its carbon factor sources are kept up to date (RSSB, 2015).

The UKWIR Carbon Accounting Workbook estimates operational carbon emissions across the UK water industry. It has been used for over ten years and is updated to reflect changes in industry needs and practices.

The National Highways Carbon Calculator is developed by National Highways, which manages and develops England's motorways and major roads. It is a spreadsheet-based calculation tool using LCA method. It divides the carbon reporting items into 11 categories including transport; bulk material; earthworks; civil, structure, and retaining walls; road pavement; drainage, fencing, barriers, and road restraint systems; street, furniture, and electrical equipment; waste; fuel, energy, and water; business and employee transport. Its carbon factors are extracted from DEFRA (Department for Environment Food & Rural Affairs) Carbon Factors 2022 & ICE (Inventory of Carbon and Energy) Carbon Factors Version 3.

The BECD is prepared by a consortium of professional bodies and organizations operating across all aspects of the UK built environment. BECD aims to align reporting practices and bring together existing carbon data in a single, consistent, free access and purportedly easy-to-use platform. The BECD has two sections: the entity level and the product level. The entity level database provides benchmark type data points to support the feasibility, early design and end of life stages. The product level database supports the evolving and detailed design, construction, and operational stages, and provide good quality product data to conduct reliable assessments. At the entity level, it will collect entity metadata (including entity details, type, and location) and project stage data (including project stage metadata, materials, energy and water, carbon emissions).

These standards, guidelines, and tools form a set of references for pursuing carbon management. They evidence good progress in the construction sector. But the lack of sophistication in data collection and reporting methods reflect the relatively experimental and early stage that the sector is in when it comes to carbon management. Taking the UK as a case, it can be observed that the construction sector has seen rapid development over the past 3-4 years with the mainstreaming of an urgent carbon reduction movement. However, there is still much to be done even just to apply existing technologies to managing the challenge. Building on all these good works, a step further would be to develop industry capability in a way that responds to carbon data trustworthiness issues that is currently holding it back.

Data trustworthiness

Data trustworthiness, which in essence relates to the ability to ascertain the correctness of the data provided by a data source (Haron et al., 2017), is a primary concern in carbon data management because informed decision-making is reliant on the availability of "good" data. Several papers explored key issues associated with data trustworthiness possible related applications for carbon data management, for example: Karthik and Ananthanarayana (2016) highlighted data reliability as a key issue for wireless sensing; Bertino et al. (2009) emphasized the data usefulness is critical for trustworthiness assurance, Haron et al. (2017) data provenance and timeliness. The limitations in practice are limiting data trustworthiness and while both researchers and practitioners working in this field are likely to be aware of this, papers that comprehensively address data trustworthiness in carbon data management are lacking.

Generally, data trustworthiness is mix-used with terminologies such as data integrity, data reliability, data quality, data representativeness, which is quite confusing. Literature on data trustworthiness is not agreed about the scope of the concept and included terminologies. Data reliability is sometimes used interchangeably with data trustworthiness (Mangel et al., 2021). According to Bertino et al. (2009), data trustworthiness includes data quality and provenance, while data integrity is a part of data quality. While Wang et al. (2011) suggests data trustworthiness as an essential parameter of assessing the data quality. Data representativeness is a significant dimension of data quality: the data is deemed of high quality if it correctly represents the real-world construct it refers to and if it fits for the intended uses in decision making and applications (Bertino et al., 2009). It is hard to argue the containment or causation relationship between data quality and data trustworthiness, quality is essential for trustworthiness while trustworthiness ensures quality.

Data trustworthiness is also associated closely with confidence about data provenance and semantic integrity, and reputation techniques (Bertino et al., 2009; Bertino, 2015). Data provenance is related to the trustworthiness of data sources and intermediaries, which can be computed using indicators such as data similarity, path similarity, data conflict, and data deduction (Dai et al., 2008). Semantic integrity concerns data consistency and correctness but it can determine whether some data correctly reflect the real world and are provided by some reliable and accurate data source (Bertino et al., 2009). Reputation techniques compute reputation scores of a system and can be used to assess data sources and data manipulation intermediaries, highly relevant to data provenance (Bertino, 2015).

Data security is also a unneglectable pillar of data trustworthiness. Only after ensuring their data is safe and their privacy is well protected, will value chain stakeholders be willing to contribute to data sharing and trust data from others. There have been various studies on data security that can be applied to carbon data management. For example, Wu et al. (2019) propose a solution for trustworthy and privacy-aware mobile crowd sensing with no need of a trusted third party to enable benign users to request tasks, contribute their data, and earn rewards anonymously without any data linkability; Abdalzaher and Muta (2020) develop a game-theoretic approach for enhancing security and data trustworthiness in IoT applications.

Based on the above review and our best knowledge, the relationship among different concepts can be summarized as shown in Figure 1. Data trustworthiness covers data availability, data quality, data compatibility, and data security. Figure 1 shows how other commonly used terms are related. Data availability and data quality have some overlap. Although data compatibility was not discussed in data trustworthiness literature, when there are different sources of data, which is usually the case in the construction sector, it becomes an essential dimension that will impact data trustworthiness.

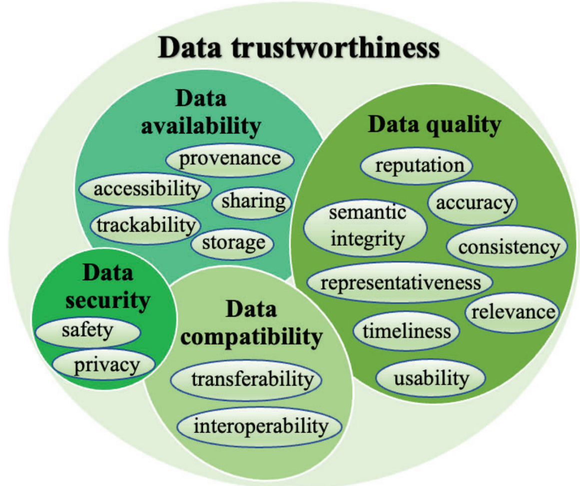

Figure 1: Data trustworthiness and related concepts

A carbon data trustworthiness framework

To address the carbon management challenges in construction projects, we propose a carbon data trustworthiness framework based on the definition of data trustworthiness, as displayed in Figure 2. This framework is based on four key pillars: data availability, data quality, data compatibility, and data security, as shown in Figure 1, and the flow between them. This framework can be used to guide carbon data model design. The principles and requirements listed in this framework provide baseline considerations for developing a trustworthy carbon data management system. They present proposed specifications for applying the framework in practice.

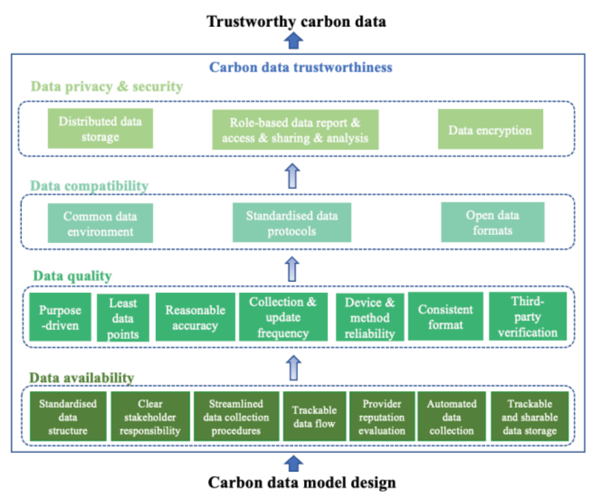

Figure 2: Carbon data trustworthiness framework

Data availability

Data availability sets the requirements about data sources. No matter whether the carbon data comes from passive reporting (as most current practices do) or from proactive monitoring, it should ideally follow a predefined standard data structure for consistency. The responsibilities of different value chain stakeholders concerning carbon data collection, accounting, and reporting should be clearly defined and agreed. A person of responsibility should be designated and trained. They should follow streamlined data collection procedures based on the construction processes to allow for data accountability. When data flows from one procedure to another, from one stakeholder to another, it should be trackable from provenance to different stages of data processing and consuming. There should be a clear data flow that every single data entry follows and a mechanism to evaluate the data providers' reputation, as a key performance indicator. The use of advanced technologies such as Internet of Things (IoT) devices to collect more detailed data (Xu et al., 2020) and data mining from other existing digital systems for automated carbon data sourcing and enhanced data availability.

Data quality

Data quality requirements are dependent on the purposes of data use. To analyze carbon management performance and identify patterns and trends for efficient carbon mitigation actions, data should be accurate, consistent, timely updated, and reliable. Firstly, when designing carbon data model, the purpose should be clarified. It is not always the case that more data is better, but choosing the right and representative data points matter. Data collection requires investment, and the cost-effective way is to be clear about what to collect, and only collect the most valuable data but with reasonable accuracy and frequency. The data accuracy, as well as data collection and update frequency, are also dependent on the purpose. The timeliness should follow the dynamics of construction works as well. The updated data shall be shared along the data flow among value chain stakeholders. Besides, attention should be paid to conduct regular examinations to ensure the reliability of data collection devices and methods. No matter the devices and methods, the data formats should be consistent to allow easy data storage, sharing, analysis, and interpretation. Finally, to validate the data quality, the use of third-party audits and certifications to verify carbon data management is also beneficial at some critical points to provide some form of assurance.

Data compatibility

To guarantee the data sharing and fusion among value chain stakeholders, carbon data should be compatible, consistent and comparable over time. The carbon data model requires interoperability among different data platforms, transferability among stakeholders, and compatibility among different devices and versions. A common data environment, open data formats, standardized data protocols are desirable. A common data environment allows for the sharing of data between different stakeholders and organizations, improving collaboration and coordination in decision-making processes. Open data formats allow for easy access to data by a wide range of users, including (possibly) the general public. This increases transparency and accountability and helps to build trust in the carbon data. Standardized data protocols help to ensure that data from different sources can be easily integrated, enabling the creation of a comprehensive and accurate picture of carbon management in construction. By standardizing the format and protocols for data, the data quality can be consistent and improved, reducing the risk of errors and increasing the reliability of the information. The use of open data formats and standardized data protocols can help to reduce the costs associated with carbon data collection, management, and analysis, as well as the costs associated with integrating data from different sources.

Data security

Data security, which includes safety and privacy, is the last but foremost dimension of carbon data trustworthiness. Carbon data relates to not only the construction activities, but also human behaviors, financial and cost details, company strategies and technologies which are business secrets. To protect the privacy and safety is critical for the collection of carbon data and the performance of carbon management. There are several technologies that can be used according to Xu et al. (2022): (1) Distributed data storage, one of the key techniques of blockchain, can allow for secured storage. It stores a copy of data at different places to avoid single point of failure. Cloud-based data platforms makes this plan feasible. (2) Role-based data reporting, access, sharing, and analysis is another significant mechanism to ensure data security. This is where only designated roles can have the right to execute designated actions to the data. This can largely avoid privacy and security violations. Finally, data encryption technologies may help with privacy protection. Blockchain is an emerging technology that is showing potential in ensuring privacy protection and data security.

Application of the framework

The four pillars, i.e., data availability, quality, compatibility, and security, and their principles can help generate trustworthy carbon data. They can work as a guidance when planning and designing carbon data model. The construction industry should start with clearly defining the roles and responsibilities, procedures of carbon management at a sector level. It would be even better to work across different sectors to ensure the compatibility across related sectors. They are encouraged to apply the framework with advanced technologies such as: building information models, sensing devices, internet of things (IoT) and computer vision for automated carbon data collection (Xu et al., 2020). Mining data from the existing data sources such as a bill of quantity, procurement database, site log or an enterprise resource platform database can also help with expanded data availability. All the data will be mapped to the carbon management system database for open but secure sharing among value chain stakeholders with the help of BIM models and blockchain. With trustworthy carbon data, the use of advanced analytics techniques, such as big data analytics, machine learning, and simulation technologies could also be applied to support carbon management decision-making such as automatic carbon calculation, carbon auditing, priority weighting, plan selection, and cost-benefit analysis in construction projects. These technologies that support the achieving of carbon data trustworthiness is displayed in Figure 3.

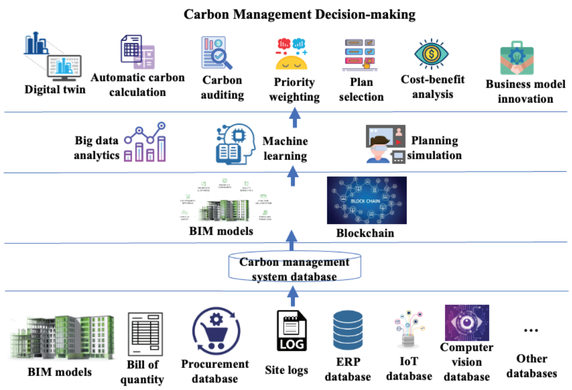

Figure 3. Technologies to support the data trustworthiness framework application

Conclusions

The construction sector is a main contributor to global carbon emissions. It is urged to take proactive measures to reduce its carbon emissions. However, obtaining reliable data remains a challenge in the construction industry. This is due to a variety of factors, including poor data collection methods, data silos, and a lack of standardization in data reporting. Without trustworthy carbon data, the current carbon data management efforts are experience-based, tedious and outputs not widely accessible or usable for relevant stakeholders.

Based on a literature review on existing carbon standards, guideline, tools, and data trustworthiness, this article clarified the definition and key pillars of data trustworthiness and developed a carbon data trustworthiness framework. The proposed framework emphasizes the significance, principles and related technologies of data availability, data quality, data compatibility, and data security. By using this framework, construction organizations can collectively improve their ability to report, track and manage carbon emissions and help to promote carbon data transparency across the sector. This will better equip them for making informed decisions to meet sustainability goals and contribute to the overall effort to combat climate change.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034337.

The authors are concurrently self-publishing a companion industry-oriented white paper on carbon data trustworthiness.

References

Abdalzaher, M. S., & Muta, O. (2020). A game-theoretic approach for enhancing security and data trustworthiness in IoT applications. *IEEE Internet of Things Journal*, *7*(11), 11250-11261.

Bertino, E. (2015). Data trustworthiness—approaches and research challenges. In *Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance: 9th International Workshop, DPM 2014, 7th International Workshop, SETOP 2014, and 3rd International Workshop, QASA 2014, Wroclaw, Poland, September 10-11, 2014. Revised Selected Papers* (pp. 17-25). Springer International Publishing.

Bertino, E., Dai, C., & Kantarcioglu, M. (2009). The challenge of assuring data trustworthiness. In *International Conference on Database Systems for Advanced Applications* (pp. 22-33). Springer, Berlin, Heidelberg.

BSI. (2023). PAS 2080: 2023 Carbon management in buildings and infrastructure. BSI Standards Limited. ISBN 9780539180343.

Cheng, M., Lu, Y., Zhu, H., & Xiao, J. (2022). Measuring CO2 emissions performance of China's construction industry: A global Malmquist index analysis. *Environmental Impact Assessment Review*, *92*, 106673.

Dai, C., Lin, D., Bertino, E., & Kantarcioglu, M. (2008). An approach to evaluate data trustworthiness based on data provenance. In *Workshop on Secure Data Management* (pp. 82-98). Springer, Berlin, Heidelberg.

Habib, S. J., & Marimuthu, P. N. (2016). Development of analytical model for data trustworthiness in sensor networks. In *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)* (pp. 467-472). IEEE.

Haron, N., Jaafar, J., Aziz, I. A., Hassan, M. H., & Shapiai, M. I. (2017). Data trustworthiness in Internet of Things: A taxonomy and future directions. *In 2017 IEEE conference on big data and analytics (ICBDA)* (pp. 25-30). IEEE.

Huang, L., Krigsvoll, G., Johansen, F., Liu, Y., & Zhang, X. (2018). Carbon emission of global construction sector. *Renewable and Sustainable Energy Reviews*, *81*, 1906-1916.

IPCC. (2018). IPCC, 2018: Annex I: Glossary [Matthews, J.B.R. (ed.)]. In: Global Warming of 1.5°C. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 541-562.

Karthik, N., & Ananthanarayana, V. S. (2016). Data trustworthiness in wireless sensor networks. In *2016 IEEE Trustcom/BigDataSE/ISPA* (pp. 2181-2186). IEEE.

Karthik, N., & Ananthanarayana, V. S. (2017). Sensor data modeling for data trustworthiness. In *2017 IEEE Trustcom/BigDataSE/ICESS* (pp. 909-916). IEEE.

Mangel, S., Gleim, L., Pennekamp, J., Wehrle, K., & Decker, S. (2021). Data Reliability and Trustworthiness through Digital Transmission Contracts. In *European Semantic Web Conference* (pp. 265-283). Springer, Cham.

One Click LCA. (2022). Construction Carbon Regulations in Europe: Review & Best Practices. Available at: https://www.oneclicklca.com/construction-carbon-regulations-in-europe/

Pouryazdan, M., Kantarci, B., Soyata, T., Foschini, L., & Song, H. (2017). Quantifying user reputation scores, data trustworthiness, and user incentives in mobile crowd-sensing. *IEEE Access*, 5, 1382-1397.

RSSB. (2015). https://safety.networkrail.co.uk/wp-content/uploads/2017/12/Rail-Carbon-Tool-User-Guide.pdf

United Nations (UN). (2015). Adoption of The Paris Agreement. Available at: https://unfccc.int/sites/default/files/english_paris_agreement.pdf (Accessed on Jan 18, 2023)

Wang, X., Govindan, K., & Mohapatra, P. (2011). Collusion-resilient quality of information evaluation based on information provenance. In *2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks* (pp. 395-403). IEEE.

Wu, H., Wang, L., Xue, G., Tang, J., & Yang, D. (2019). Enabling data trustworthiness and user privacy in mobile crowdsensing. *IEEE/ACM Transactions on Networking*, *27*(6), 2294-2307.

Xu, J., Chen, K., Zetkulic, A. E., Xue, F., Lu, W., & Niu, Y. (2020). Pervasive sensing technologies for facility management: A critical review. *Facilities*, *38*(1/2), 161-180.

Xu, J., Lu, W., Wu, L., Lou, J., & Li, X. (2022). Balancing privacy and occupational safety and health in construction: A blockchain-enabled P-OSH deployment framework. *Safety science*, *154*, 105860.